AD-A244 042

N° d'ordre:

(2)

THESE

DAAL03-90-C-0029

présentée à

L'UNIVERSITÉ DE PROVENCE
CENTRE ST-CHARLES

DTIC
ELECTE
DEC 7 1991
S C D

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE PROVENCE

Spécialité :

MATHÉMATIQUES APPLIQUÉES

par

Marie-Christine ROUBAUD

Sujet de la thèse :

FILTRAGE LINÉAIRE PAR MORCEAUX AVEC PETIT BRUIT D'OBSERVATION

Soutenue le 19 Novembre 1990 devant le jury composé de :

Mme	**Jacqueline DETRAZ**	Présidente
Mme	**Dominique MICHEL**	Examinateurs
MM	**Peter CAINES**	
	Etienne PARDOUX	
	Jean PICARD	

91-17367

91 1209 033

N° d'ordre:

Accession For
NTIS GRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By per Form 50
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

THESE

présentée à

L'UNIVERSITÉ DE PROVENCE
CENTRE ST-CHARLES

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE PROVENCE

Spécialité :

MATHÉMATIQUES APPLIQUÉES

CLASSE MATHEMATIQUE 4

par

Marie-Christine ROUBAUD

Sujet de la thèse :

(Piecewise linear Filtering with Small observation Noise)

FILTRAGE LINÉAIRE PAR MORCEAUX AVEC PETIT BRUIT D'OBSERVATION

Soutenue le 19 Novembre 1990 devant le jury composé de :

Mme	**Jacqueline DETRAZ**	Présidente
Mme	**Dominique MICHEL**	Examinateurs
MM	**Peter CAINES**	
	Etienne PARDOUX	
	Jean PICARD	

Je remercie tout d'abord Etienne Pardoux, le directeur de cette thèse pour les idées et les conseils qu'il m'a donnés dans l'élaboration de ce travail, ainsi que pour son dynamisme et son efficacité.

Je tiens à exprimer tous mes remerciements à Dominique Michel et Jean Picard d'avoir accepté d'être rapporteurs et membres du jury de cette thèse ainsi qu'à Jacqueline Detraz de me faire l'honneur de le présider et Peter Caines d'y participer .

Je remercie l'équipe du projet MEFISTO à l'INRIA pour leur accueil chaleureux et pour m'avoir permis de terminer ce travail dans de très bonnes conditions.

Je remercie également mes collègues du Laboratoire de Mathématiques Appliquées et Informatique de l'Université de Provence qui par leur sympathie ont rendu l'ambiance de travail très agréable et furent d'un grand soutien durant ces années.

Je tiens à remercier tout particulièrement Fabien Campillo, François Le Gland et Catherine Donati pour leur disponibilité et leur aide souvent sollicitées.

Enfin, j'adresse tous mes remerciements à ma famille qui a su me soulager de bien de tâches matérielles dans les moments difficiles et où j'ai puisé l'énergie nécessaire pour réaliser ce document.

Introduction

Le problème de filtrage consiste à estimer à chaque instant t, une fonctionnelle $\varphi(X_t)$ d'un processus non observé $\{X_t\,;\,t \geq 0\}$ au vu des réalisations jusqu'à l'instant t du processus observé $\{Y_t\,;\,t \geq 0\}$. L'estimateur optimal au sens de la moyenne quadratique étant l'espérance conditionnelle $\mathbf{E}(\varphi(X_t)|\mathcal{Y}_t)$ avec $\mathcal{Y}_t$ la tribu des observations jusqu'à l'instant t, on cherche donc à caractériser la loi conditionnelle de X_t sachant $\mathcal{Y}_t$.

Pour un système linéaire gaussien, il est bien connu que la loi conditionnelle est gaussienne et que sa moyenne et sa variance sont respectivement solution d'une équation différentielle stochastique linéaire et d'une équation de Riccati. Ce filtre a été proposé dans les années soixantes par R.E. Kalman et R.S. Bucy [16]. Pour les besoins des applications pratiques, la théorie du filtrage linéaire a été étendue sans justification mathématique rigoureuse aux systèmes non linéaires et le filtre de Kalman étendu (cf.[14]) est l'algorithme le plus fréquemment utilisé par les ingénieurs notamment dans le domaine de l'aérospatiale. Cependant, ses résultats peuvent être très mauvais. Dans [33], on trouve des exemples où le filtre de Kalman étendu est inefficace.

Dans le domaine théorique, des équations aux dérivées partielles stochastiques décrivant l'évolution de la loi conditionnelle et de sa version non normalisée lorsqu'elle existe, ont été développées par de nombreux auteurs (cf.[20], [38], [11], [23]). Mais en général, la résolution numérique de ces équations pose de grosses difficultés, particulièrement pour les systèmes de grandes dimensions. D'autre part, des travaux récents (voir par exemple [4], [27]) ont mis en évidence qu'il y a très peu de situations où la loi conditionnnelle peut être caractérisée par un nombre fini de statistiques récursives. Donc, il est intéressant d'obtenir des filtres approchés de dimension finie, facilement calculables et d'étudier des situations pour lesquelles le problème de filtrage non linéaire peut être résolu asymptotiquement.

Ces dernières années, les problèmes asymptotiques lorsque le bruit d'observation tend vers zéro ont suscité beaucoup d'intérêt. Lorsque la fonction d'observation est injective, sous certaines hypothèses de régularité, il a été montré que la variance conditionnelle est asymptotiquement nulle et plusieurs filtres approchés ont été proposés (cf. [18], [30]-[32], [2], [15]). Dans ce contexte, J. Picard dans [31] et D. Ji dans [15], ont prouvé que le filtre de Kalman étendu était efficace. Ceci ne sera plus le cas dans les situations que nous nous proposons d'étudier.

Dans cette thèse, nous nous intéressons à des problèmes de filtrage "linéaire par morceaux" avec petit bruit d'observation dans le cas où la fonction d'observation est non injective. Notre étude s'inscrit dans le cadre général où le processus à estimer $\{X_t\,;\, t \geq 0\}$ à valeurs dans $\mathbb{R}^n$ est une diffusion dont le coefficient de dérive est continu et linéaire par morceaux et le coefficient de diffusion est constant par morceaux. On observe le processus $\{Y_t\,;\, t \geq 0\}$ à valeurs dans $\mathbb{R}^d$, défini par

$$Y_t = \int_0^t h(X_s)ds + \varepsilon W_t,$$

où la fonction d'observation h est continue et linéaire par morceaux, $\{W_t\,;\, t \geq 0\}$ est un processus de Wiener non nécessairement indépendant du signal et ε est un paramètre supposé petit. Si nous désignons par $\{\Theta_i\}_{i=1,l}$ la partition polyédrale de $\mathbb{R}^n$ sur laquelle est défini le problème lineaire par morceaux, à chaque Θ_i, on peut associer un problème linéaire $(\mathcal{P}_i)$ avec condition initiale non gaussienne.

En exploitant le fait que l'observation est faiblement bruitée et la linéarité locale des systèmes considérés, sous certaines "hypothèses de détectabilité", nous proposons des filtres sous-optimaux facilement calculables et nous obtenons des résultats asymptotiques lorsque ε tend vers zéro. Nous travaillons sur un intervalle de temps $[0, T]$ fixé et de manière générale, nos estimations ne sont pas uniforme en t.

Dans le cas d'un processus d'état unidimensionel avec un coefficient de diffusion constant et sous l'hypothèse d'indépendance des bruits d'évolution et d'observation, W.H. Fleming, D. Ji et E. Pardoux ont montré dans [7] que sous une certaine "hypothèse de détectabilité" que nous noterons $(HD1)$, on peut construire un filtre approché de dimension finie qui utilise une batterie de l filtres de Kalman-Bucy et une procédure de deux tests statistiques. Un premier test permet de détecter les intervalles de temps sur lesquels une

trajectoire $t \to X_t$ prend ses valeurs loin des points critiques de la fonction h et sur ces intervalles, un second test permet de décider lequel des l filtres de Kalman-Bucy suivre.

Cet algorithme repose sur l'idée suivante: durant les intervalles de temps sur lesquels une trajectoire $t \to X_t$ prend ses valeurs loin des points critiques, si la variance conditionnelle est petite, la densité de la loi conditionnelle n'a qu'un seul pic significatif et elle s'approche d'une densité gaussienne. L'hypothèse $(HD1)$ est une condition suffisante pour que cette situation soit vérifiée. Nous remarquons que cette hypothèse n'intervient pas dans la détection des intervalles de monotonie de la fonction h. Dans [9], une approche similaire a été développée par W.H. Fleming et E. Pardoux dans le cas d'une fonction d'observation monotone par morceaux. Cette idée de base constitue le point de départ de notre travail.

Cette thèse se compose de trois chapitres qui peuvent être lus indépendamment.

Dans le chapitre I, on traite le cas dim X = dim $Y \geq 1$. Notre motivation principale est de considérer lorsque l'hypothèse $(HD1)$ n'est pas vérifiée, une autre situation où la variance conditionnelle est faible. On étudie en détail le cas $l = 2$. Sous des hypothèses dites du type (2), nous proposons le même type d'approximations du filtre optimal que dans [7]. Sur les intervalles de linéarité de la fonction h, les tests statistiques permettant de décider sur quel polyèdre de la partition la trajectoire $t \to X_t$ prend ses valeurs, sont de la forme rapport de vraisemblance. Ils sont construits à l'aide du théorème de Girsanov.

D'autre part, nous étendons les travaux de W.H. Fleming, D. Ji et E. Pardoux [7] a une équation de dynamique linéaire par morceaux plus générale et dans le cas dim X = dim $Y > 1$. Ceci nous conduira à introduire des hypothèses dites du type (1) incluant $(HD1)$. La principale difficulté de cette extension réside dans le fait que le coefficient de diffusion n'est plus supposé constant mais constant par morceaux. Nous ne pouvons plus par un simple changement de probabilité écrire le problème de départ sous forme linéaire. Nous devons introduire un processus d'état intermédiaire avec un coefficient de diffusion constant.

Les erreurs d'approximation du filtre optimal sont estimées. La différence majeure entre les deux types de situations étudiées est que lorsque ε tend vers zéro, sous les hypotheses du type (1), les probabilités d'erreur des tests

convergent vers zéro tandis que sous les hypothèses du type (2), elles convergent vers une quantité dépendante de la longueur de l'intervalle de temps sur lequel est prise une décision. Dans le second cas, les tests ne décident correctement que sur des intervalles de temps suffisament longs.

Les résultats obtenus ont été étendus au cas $l \geq 2$ en supposant le coefficient de diffusion constant. Par sa construction, le test de détection des passages de $t \rightarrow X_t$ par les frontières des polyèdres Θ_t devient très pénalisant dès que l est grand. Dans ce cas, ces procédures ne sont intéressantes que lorsque les valeurs prises par la fonction h autorisent le découpage de leurs domaines d'application.

Ceci constitue une limitation à l'utilisation de nos résultats pour la construction de filtres sous-optimaux pour un problème de filtrage non linéaire quelconque en approchant les coefficients par des fonctions linéaires par morceaux. Par ailleurs, nos résultats peuvent s'adapter à des problèmes de filtrage non linéaire avec une fonction d'observation injective par morceaux (cf.[9])

Dans le chapitre II, les algorithmes présentés dans le chapitre précédent ont été mis en œuvre dans le cas particulier où $l = 2$, dim X = dim $Y = 1$ et en supposant le bruit d'observation indépendant du signal. Nous adaptons notre étude en temps continu au temps discret. Ce travail a été fait en collaboration avec P. Milheiro de Oliveira.

En dimension un, les deux types de situations considérées se résument à deux hypothèses $(HD1)$ et $(HD2)$. Nous nous intéressons principalement à la situation où $(HD2)$ est vérifiée, la discrétisation du problème sous $(HD1)$ ayant déja été traitée dans [8]. Les formules explicites permettant de calculer les bornes des différents tests ainsi que les temps moyens d'attente pour prendre une décision sont justifiées par un raisonnement heuristique en approchant comme dans [8], certains processus discrets par des diffusions.

Nous étudions sur divers exemples le comportement des filtres proposés par rapport à ceux du processus à estimer et du filtre optimal obtenu de façon approchée par résolution de l'équation de Zakai discrétisée (cf. [22]). Nous comparons les différentes procédures permettant d'obtenir ces filtres approchés et étudions l'influence de certains paramètres du système sur leurs performances. On vérifie notamment que les temps moyens d'attente théoriques et empiriques pour prendre une décision sont plus grands sous l'hypothèse $(HD2)$ que sous $(HD1)$.

Dans le chapitre III, nous abordons le cas dim $X >$ dim Y en étudiant une classe particulière de problèmes. Dans ce contexte, pour adapter les idées développées précédemment, nous devons trouver des situations pour lesquelles la variance conditionnelle est petite selon les directions de découpage de la partition.

Dans le problème linéaire partiellement observé avec petit bruit d'observation, S.S. Sachs montre dans [34] qu'uniquement les composantes directement mesurées du processus d'état ont une variance conditionnelle asymptotiquement nulle lorsque ε tend vers zéro. Dans notre optique, ne sont donc intéressants que les problèmes où les composantes de X_t correspondantes aux directions de découpage sont directement mesurées dans les problèmes linéaires $(\mathcal{P}_i)$ associés à la partition.

Dans le cas traité, on suppose dim $X = 2$, dim $Y = 1$, $l = 2$ et on prend pour direction de découpage le vecteur $(1,0)^* \in \mathbb{R}^2$. Le processus de diffusion $\{(X_1(t), X_2(t))^* ; t \geq 0\}$ à estimer, a un coefficient de dérive linéaire par morceaux et un coefficient de diffusion constant. Il est observé par un processus unidimensionnel faiblement bruité défini par

$$Y_t = \int_0^t h_1(X_1(s))ds + \varepsilon W_t ,$$

où la fonction h_1 est supposé non injective et le processus de Wiener $\{W_t ; t \geq 0\}$ indépendant du signal. La non linéarité du système dépend uniquement de la composante du processus d'état pouvant être estimée de façon précise.

L'algorithme décrit dans le chapitre I sous l'hypothèse $(HD1)$, permet de proposer sur l'intervalle de temps $[0,T]$, un filtre approché $\bar{X}_1(t)$ pour la composante $X_1(t)$. Il n'en est pas de même pour $X_2(t)$. Dans le cas linéaire, la composante du filtre de Kalman-Bucy qui lui correspond n'est pas à "mémoire courte" comme dans le chapitre I, l'estimation de $X_2(t)$ nécessite la connaissance de toute l'histoire de l'observation jusqu'à l'instant t. En s'inspirant du cas $\varepsilon = 0$ et du fait que sous les hypothèses émises, la loi conditionnelle de $X_2(t)$ sachant $\sigma(X_1(s) ; 0 \leq s \leq t)$ est gaussienne, on construit à partir de $\bar{X}_1(t)$ un filtre approché pour la composante $X_2(t)$.

Lorsque $\varepsilon \to 0$, on montre que l'erreur d'approximation de $X_1(t)$ est asymptotiquement nulle et que celle de $X_2(t)$ est asymptotiquement optimale au sens où elle converge vers l'erreur de filtrage du processus $\{X_2(t) ; t \geq 0\}$ observé par $\{X_1(t) ; t \geq 0\}$.

Tout au long de ce travail, l'hypothèse d'un grand rapport signal/bruit

est fondamentale. Lorsque cette hypothèse n'est pas vérifiée, une approche différente du problème s'impose. Nous signalons que C. Savona a proposé dans [35], un autre type d'algorithme pour obtenir des filtres sous–optimaux en discrétisant le temps et en exploitant le caractère linéaire par morceaux des coefficients dans le cas dim X = dim Y.

Bibliographie

[1] R.F. Bass, E. Pardoux : Uniqueness for diffusions with piecewise constant coefficients, in *Probability Theory Rel. Fields* Springer-Verlag, **76**, 557-572, 1987.

[2] A. Bensoussan : On some approximation techniques in non linear filtering, in *Stoch. Diff. Syst., Stoch. Control and Appl.* W.H. Flemming & P.L. Lions eds., IMA Vol **10**, 17-31, Springer, 1987.

[3] P. Billingsley : Convergence of Probability Measures, John Wiley & Sons, New York, 1968.

[4] M. Chaleyat-Maurel, D. Michel : Des résultats de non existence de filtre de dimension finie, *Sthocastics*, **13**, 83-102, 1984.

[5] N. El Karoui, M. Chaleyat-Maurel : Un problème de réflexion et ses applications au temps local et aux équations différentielles stochastiques sur $\mathbb{R}$, cas continu, *Astérisque* , **52-53**, 117-144, 1978.

[6] R.S. Ellis : Entropy, Large Deviations, and Statistical Mechanics. Springer-Verlag, 1985.

[7] W.H. Fleming, D. Ji, E. Pardoux : Piecewise linear filtering with small observation noise, in *Analysis and Optimization of Systems*. Lecture Note in Cont. and Info. Sci. **111**, Springer, 725-729, 1988.

[8] W.H. Fleming, D. Ji, P. Salame et Q. Zhang : Discrete time piecewise linear filtering with small observation noise, *Brown Univ. LCDS/CCS* report **N88-27** , 1988.

[9] W.H. Fleming, E. Pardoux : Piecewise monotone filtering with small observation noise, *Siam J. Control Optim.* **20**, 261-285, 1989.

[10] W.H. Fleming and R.W. Rishel : Deterministic and Stochastic Optimal Control, Tome 1, Applications of Mathematics, Springer-Verlag, 1975.

[11] M. Fujisaki, G. Kallianpur, and H. Kunita : Stochastic differential equations for the nonlinear filtring problem, *Osaka J. Math.* **9**, 19-44, 1972.

[12] A.H. Haddad : Linear filtering of singularly perturbed systems, *IEEE Transactions on Automatic Control*, **AC-21**, 515-519, 1976.

[13] U. Haussmann, E. Pardoux : A conditionnally almost linear filtering problem with non Gaussian initial condition, *Stochastics*, **23**, 241-275, 1988.

[14] N. Ikeda, S. Watanabe : Stochastic Differential equations and diffusions. North-Holland, 1981.

[15] A.H. Jazwinski : Stochastic Processes and Filtering Theory, Academic Press, 1970.

[16] D. Ji : Nonlinear filtering with small observation noise, *Ph. D. Thesis*, Brown University, 1987.

[17] R.E. Kalman, R.S. Bucy : New results in linear filtering and prediction theory, *Journal of Basic Engineering*, 95-108, 1961.

[18] I. Karatzas, S.E. Shreve : Brownian Motion and Stochastic Calculus, Springer-Verlag, 1988.

[19] R. Katzur, B.Z. Bobrovsky and Z. Schuss : Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel, *SIAM J Appl. Math.*, **44**, 1984, Part I: 591-604, Part II: 1176-1191.

[20] A.J. Krener : The asymptotic approximation of non linear filters by linear filters, *in Proc. 7th MTNS Symposium*, Stockholm, 1985.

[21] H.J. Kushner : On the differential equations satisfied by the conditional probability densities of Markov processes, *SIAM J. Control*, **2**, 106-119, 1964.

[22] H. Kwakernaak, R. Sivan : The Maximally achievable accuracy of linear optimal regulators and linear optimal filters. *IEEE Transactions on Automatic Control*, **AC-16**, 79-86, 1972.

[23] F. Le Gland : Estimation de paramètres dans les processus stochastiques, en observation incomplète – Application à un problème de radioastronomie, *Thèse de Docteur-Ingenieur*, Université Paris IX-Dauphine, 1981.

[24] R.CH. Liptzer, A.N. Shyriaev : Statistics of Random Processes I-II. Springer-Verlag, 1977.

[25] P. Milheiro de Oliveira : Filtres approchés pour un problème de filtrage non linéaire discret avec petit bruit d'observation,*rapport INRIA*, **1142**. 1989.

[26] P. Milheiro de Oliveira : Etudes asymptotiques en filtrage non linéaire avec petit bruit d'observation, *Thèse*, Université de Provence, septembre 1990.

[27] P.J. Moylan : A note on Kalma-Bucy filters with zero measurement noise, *IEEE Trans. Inform. Th.* **19**, 263-264, 1974.

[28] D. Ocone, E. Pardoux : A Lie-algebraic criterion for non existence of finite dimentionnally computable filters. *in Stochastic Partial Differential Equations and Applications II*, G. Da Prato and L. Tubaro eds., Lectures Notes in Math., **1390**, 1989.

[29] E. Pardoux, M.C. Roubaud : General piecewise linear filtering problems with small observation noise, *in* 28^{th} *IEEE CDC*, Tampa, 1989.

[30] E. Pardoux, C. Savona : Piecewise linear filtering, *in Stochastic Differential Systems, Stochastic Control Theory and Appl.*, W. Fleming and P.L. Lions eds., IMA volume in Mathematics and its Appl. **10**, Springer Verlag, 1987.

[31] J. Picard : Non linear filtering of one-dimensional diffusions in the case of a high signal -to-noise ratio, *Siam J Apl. Math.* **46**, 1098-1125,1986.

[32] J. Picard : Non linear filtering and smoothing with high signal-to-noise ratio, *Stochastic Processes in Physics and Engineering*, (Bielefeld 1986), Reidel, 237-251, 1988.

[33] J. Picard : Asymptotic study of estimation problems with small observation noise, *in Stochastic Modelling et Filtering*, (Rome 1984), *Lect. Notes in Control and Inf. Sci.*, **91**, Springer, 1987.

[34] J. Picard : Efficiency of the extended Kalman filter for non linear systems with small noise, *Rapport INRIA* , **1068**, 1989.

[35] S.S. Sachs : Asymptotic analysis of linear filtering problems, *M.S. Thesis*, Departement of Systems Engineering, Case Western Reserve University, Cleveland, Ohio, June 1980.

[36] C. Savona : Approximate nonlinear filtering for piecewise linear systems, *Systems & Control Letters, North-Holland*, **11**, 327-332, 1988.

[37] W.Murray Wonham : Linear Multivariable Control: a Geometric Approach, Springer-Verlag, 1979.

[38] I. Yaesh, B.Z. Bobrovsky, Z. Schuss : Asymptotic analysis of the optimal filtering problem for two dimensional diffusions measured in a low noise channel, *SIAM Journal of Appl. Math.*, **Vol.50**, N4, pp. 1134-1155, 1990.

[39] M. Zakai : On the optimal filtering of diffusion processes, *Z. Wahrsh. Verw. Geb.*, **11**, 230-243, 1969.

Sommaire

Chapitre I : Filtrage linéaire par morceaux avec petit bruit d'observation

Chapitre II : Filtrage linéaire par morceaux d'un système en temps discret avec petit bruit d'observation

en collaboration avec P. Milheiro de Oliveira

Chapitre III : Un problème de filtrage linéaire par morceaux partiellement observé avec petit bruit d'observation

Chapitre I

Filtrage linéaire par morceaux avec petit bruit d'observation

Table des matières

I Filtrage linéaire par morceaux avec petit bruit d'observation I-1

1 Introduction I-3

2 Formalisation du problème. Lemmes préliminaires I-7

3 Détection des passages de $\{X_t\}$ par Δ_u I-20

4 Sous $(HD1)$, le test de la variation quadratique I-32

5 Sous $(HD2)$, le test du rapport de vraisemblance sur les accroissements de l'observation I-36

5.1 Cas $\varepsilon = 0$ I-36

5.2 Cas $\varepsilon > 0$ I-40

6 Le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy I-51

7 Erreurs d'approximation I-72

8 Extension des résultats au cas $l \geq 2$ I-79

Résumé

On considère un problème de filtrage linéaire par morceaux avec petit bruit d'observation dans le cas où la fonction d'observation est non injective. Dans deux types de situations différentes, on construit un filtre approché de dimension finie, à partir de plusieurs filtres de Kalman-Bucy calculés en paralléle et d'une procédure de tests. Dans le premier cas étudié, nos résultats étendent ceux de Fleming, Ji, Pardoux [4] à une équation de dynamique linéaire par morceaux plus générale.

1 Introduction

On considére un cas particulier du problème de filtrage non linéaire, le problème dit linéaire par morceaux avec petit bruit d'observation. En général, le filtre optimal étant de dimension infinie, il est intéressant de l'approcher par un filtre de dimension finie et d'obtenir des résultats asymptotiques quand le bruit d'observation tend vers zéro. C'est ce que nous nous proposons de faire pour le problème suivant:

$$\begin{cases} X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dU_s, \\ Y_t &= \int_0^t h(X_s)ds + \varepsilon W_t, \end{cases}$$

où $\{X_t, t \geq 0\}$ est le processus non observé , à estimer à l'instant t, connaissant les trajectoires jusqu'à l'instant t du processus $\{Y_t, t \geq 0\}$. Les processus $\{U_t, t \geq 0\}$ et $\{W_t, t \geq 0\}$ sont deux processus de Wiener non nécessairement indépendants et le paramètre ε est petit.

Nous étudions le cas où les processus $\{X_t, t \geq 0\}$ et $\{Y_t, t \geq 0\}$ sont n-dimensionels. Nous supposons qu'il existe une partition polyédrale finie de $\mathbb{R}^n$, $\mathbb{R}^n = \cup_{i=1}^{l} \Theta_i$, telle que sur chaque polyèdre Θ_i, les restrictions des applications continues b et h sont affines, et celles de l'application mesurable σ constantes. Sur chaque polyèdre Θ_i, le problème s'écrit comme un problème de filtrage linéaire $(\mathcal{P}_i)$ avec condition initiale non gaussienne.

Durant un certain intervalle de temps, si nous savons que la trajectoire $t \rightarrow X_t$ prend ses valeurs sur un et un seul polyèdre Θ_i, il semble naturel d'approcher le filtre optimal par le filtre de Kalman-Bucy (FK_i) correspondant au problème $(\mathcal{P}_i)$. Dans le cadre de cette étude, le filtre (FK_i) étant à "mémoire courte" quand ε tend vers zéro, la condition initiale est sans importance. Nous devons donc construire deux tests: l'un permettant de détecter des intervalles de temps durant lesquels une trajectoire $t \rightarrow X_t$ prend ses valeurs sur un et un seul polyèdre Θ_i, $i = 1, \cdots, l$, l'autre permettant de déterminer lequel. Ce deuxième type de test ne peut être construit que dans les situations où la variance de la loi conditionelle est petite tout au moins dans certaines directions.

Pour ε petit, on peut déduire de l'observation, une approximation de $h(X_t)$. Si l'application h est injective, on obtient donc une estimation de X_t. Les résultats de J. Picard [12]-[14] montrent que dans ce cas et sous certaines hypothèses de régularité, la variance de la loi conditionnelle est petite et que

le filtre de Kalman étendu est une "bonne" estimation. On s'intéresse à la situation où la fonction d'observation h est non injective. Dans ce cas, la variance conditionnelle peut être grande et nous n'obtenons une bonne approximation que sous deux types d'hypothèses.

Supposons que $h(\Theta_i) \cap h(\Theta_j) \neq \emptyset$. Les hypothèses dites de type (1) permettent une discrimination des systèmes de filtrage linéaire $\mathcal{P}_i$ et $\mathcal{P}_j$ par les parties martingales des systèmes. Sous les hypothèses dites de type (2), la différence est mise en évidence par les parties à variations bornées.

Dans le cas unidimensionel, les deux types de situations considérées se résument en deux hypothèses. Nous les présentons dans le cas simple où $\varepsilon = 0$ et les coefficients s'écrivent:

$$\begin{aligned} b(x) &= B_- x \mathbf{1}_{\{x<0\}} + B_+ x \mathbf{1}_{\{x\geq 0\}}, \\ \sigma(x) &= \Sigma_- \mathbf{1}_{\{x<0\}} + \Sigma_+ \mathbf{1}_{\{x\geq 0\}}, \\ h(x) &= H_- x \mathbf{1}_{\{x<0\}} + H_+ x \mathbf{1}_{\{x\geq 0\}}, \end{aligned}$$

avec $H_- H_+ < 0$. Dans ce cas, $h(X_t)$ est observé exactement, mais il n'est pas toujours possible de déterminer le signe de X_t ([1]). Nous savons le déterminer sous deux hypothèses:

$(HD1)$ $\qquad H_-^2 \Sigma_-^2 \neq H_+^2 \Sigma_+^2$.

Dans ce cas, la variation quadratique du processus $h(X_t)$ nous indique le signe de X_t.

$(HD2)$ $\qquad H_-^2 \Sigma_-^2 = H_+^2 \Sigma_+^2$ et $B_- \neq B_+$.

Dans ce second cas, la variation quadratique de $h(X_t)$ ne donne aucune information sur le signe de X_t. Néanmoins, si nous supposons que $\frac{dY_t}{dt} \neq 0$ pour $t \in [a,b]$, nous savons que soit $X_t > 0$ pour tout $t \in [a,b]$, soit $X_t < 0$, pour tout $t \in [a,b]$. Posons $y_t = h(X_t)$ et $\Gamma = |H_+ \sigma_+| = |H_- \sigma_-|$, nous avons:

(+) $y_t = y_a + \int_a^t B_+ y_s ds + \Gamma(U_t^+ - U_a^+)$,
où dans le cas $\{X_t > 0\,,\, t \in [a,b]\}$, le processus $\{U_t^+ - U_a^+\,,\, a \leq t \leq b\}$ est un processus de Wiener standard.

(-) $y_t = y_a + \int_a^t B_- y_s ds + \Gamma(U_t^- - U_a^-)$,
où sur $\{X_t < 0\,,\, t \in [a,b]\}$, le processus $\{U_t^- - U_a^-\,,\, a \leq t \leq b\}$ est un processus de Wiener standard.

[1] par exemple lorsque $h(x) = |x|$, $b \equiv 0$, $\sigma \equiv 1$.

Vu que $B_- \neq B_+$, un test du rapport de vraisemblance nous permet de décider entre les deux alternatives. La statistique de test s'écrit:

$$\frac{B_+ - B_-}{\Gamma^2} \int_a^b y_s dy_s - \frac{B_+^2 - B_-^2}{2\Gamma^2} \int_a^b y_s^2 ds.$$

On note que la différence majeure entre la situation sous $(HD1)$ et celle sous $(HD2)$ est que sous $(HD1)$ une décision correcte peut être prise instantanément (tout au moins théoriquement) sur l'événement $\{y_t \neq 0\,,\, a \leq t \leq a+\delta\}$, $\delta > 0$, tandis que sous $(HD2)$, la probabilité de prendre une décision correcte tend vers un quand la longueur de l'intervalle d'observation sur lequel $y_t \neq 0$ tend vers l'infini!

Dans le cadre général, nous introduirons deux hypothèses semblables à celles décrites ci-dessus (cf. paragraphes 4-5), pour construire respectivement sous $(HD1)$ et sous $(HD2)$, un test de choix sur les accroissements de l'observation. Nous considérerons deux nouvelles hypothèses respectivement du type (1) et (2), $(\overline{HD}1)$ et $(\overline{HD}2)$ (cf. paragraphe 6), sous lesquelles nous proposons un test de choix sur les sorties des filtres de Kalman-Bucy. Comme dans l'exemple précédent, la qualité du test de choix ne sera pas la même dans les deux types de situations. Il faut beaucoup plus de temps pour prendre une décision correcte sous le second type d'hypothèses que sous le premier.

Dans le cas $n = 1$, $\sigma \equiv 1$ et avec l'hypothèse d'indépendance des bruits d'évolution et d'observation, notre problème a déjà été considéré sous $(HD1)$ par W.H. Fleming, D. Ji et E. Pardoux [4]. Sous une adaptation de l'hypothèse $(HD1)$ au cas où h est monotone par morceaux, le même type de problème a été traité dans [6]. Dans [4], les passages de X_t par les points critiques de la fonction d'observation h sont détectés, avec un probabilité proche de un, par un test sur les accroissements de l'observation. Dans notre étude, nous montrons que comme dans [6], où la fonction h est régulière, un test peut être construit sur la sortie d'un des filtres de Kalman-Bucy. Le fait de considérer une équation de dynamique linéaire par morceaux plus générale que dans [4] dans le cas $n > 1$, nous conduira à démontrer certains résultats différemment et à introduire une autre hypothèse du type (1) incluant l'hypothèse $(HD1)$. Cependant la principale nouveauté de notre propos se situe dans la considération du problème sous des hypothèses du type (2).

Ce chapitre est organisé comme suit:

Dans le paragraphe 2, nous formalisons le problème, posons les hypothèses et présentons des résultats préliminaires. Par la suite, nous étudions en détail le cas $l = 2$. Dans le paragraphe 3, nous proposons un test sur la sortie d'un des deux filtres de Kalman-Bucy qui nous assure, avec une probabilité proche de un, que la trajectoire $t \to X_t$ prend ses valeurs sur un et un seul polyèdre Θ_i durant l'intervalle de temps $[a, b]$. Il reste à savoir lequel. C'est ce que nous nous proposons de déterminer dans les paragraphes 4, 5 et 6. Dans le paragraphe 4, sous l'hypothèse $(HD1)$, nous présentons un test basé sur la variation quadratique des accroissements des observations. Dans le paragraphe 5, sous l'hypothèse $(HD2)$, nous proposons un test du type rapport de vraisemblance sur les accroissements des observations. Sous $(\overline{HD}1)$ et $(\overline{HD}2)$, un second test du type rapport de vraisemblance sur les sorties des deux filtres de Kalman est donné dans le paragraphe 6. Dans le paragraphe 7, nous estimons l'erreur d'approximation du filtre optimal et finalement, en supposant le coefficient de diffusion constant, nous étendons nos résultats au cas $l \geq 2$ dans le paragraphe 8.

2 Formalisation du problème. Lemmes préliminaires

On considère le système différentiel stochastique suivant:

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dU_s, \tag{2.1}$$

$$Y_t = \int_0^t h(X_s)ds + \varepsilon(\Upsilon U_t + \bar{\Upsilon}\bar{U}_t), \tag{2.2}$$

où $\{U_t, t \geq 0\}$ et $\{\bar{U}_t, t \geq 0\}$ sont deux processus de Wiener standards indépendants à valeurs dans $\mathbb{R}^n$ et définis sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ que nous préciserons ci-dessous, X_0 est une variable aléatoire $\mathcal{F}_0$- mesurable de loi Π_0, b et h sont deux applications mesurables de $\mathbb{R}^n$ dans $\mathbb{R}^n$, σ est une application mesurable de $\mathbb{R}^n$ dans $\mathbb{R}^n \otimes \mathbb{R}^n$, Υ et $\bar{\Upsilon}$ sont des matrices (n,n) et ε est un paramètre réel positif. Deux cas sont considérés : $\varepsilon = 0$ et $\varepsilon > 0$.

Hypothèses: On suppose que

$(H1)$ X_0 est une variable aléatoire indépendante de $\{(U_t, \bar{U}_t), t \geq 0\}$ et il existe $c_0 > 0$ tel que

$$\mathbf{E}(\exp c_0|X_0|^2) < +\infty.$$

$(H2)$ Les fonctions b, h appartiennent à $C(\mathbb{R}^n; \mathbb{R}^n)$ et il existe une partition polyédrale finie de $\mathbb{R}^n$, $\mathbb{R}^n = \cup_{i=1}^l \Theta_i$, des matrices (n,n) B_i, H_i, Σ_i, $1 \leq i \leq l$ et des vecteurs b_i, $h_i \in \mathbb{R}^n$ tels que :

$$\begin{aligned} b(x) &= \textstyle\sum_{i=1}^l (B_i x + b_i)\mathbf{1}_{\Theta_i}(x) \\ \sigma(x) &= \textstyle\sum_{i=1}^l \Sigma_i \mathbf{1}_{\Theta_i}(x) \\ h(x) &= \textstyle\sum_{i=1}^l (H_i x + h_i)\mathbf{1}_{\Theta_i}(x). \end{aligned}$$

$(H3)$ la fonction h est non globalement injective.

$(H4)$ H_i est inversible , $\Sigma_i \Sigma_i^*$ est définie positive , $1 \leq i \leq l$.

$(H5)$ $\Upsilon\Upsilon^* + \bar{\Upsilon}\bar{\Upsilon}^* = I_n$ et $\Upsilon\bar{\Upsilon}^*$ est définie positive.

On pose $\Omega_1 = \Omega_2 = C(\mathbb{R}^+; \mathbb{R}^n)$. Soit $\Omega = \Omega_1 \times \Omega_2$ d'élément générique $\omega = (\omega_1, \omega_2)$, $\mathcal{F}$ sa tribu borélienne, $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, $\{(X_t, \bar{U}_t)^*, t \geq 0\}$ le processus canonique sur Ω:

$$(X_t, \bar{U}_t)(\omega_1, \omega_2) = (\omega_1(t), \omega_2(t)),$$

et $(\mathcal{F}_t)_{t\geq 0}$ la filtration naturelle du processus canonique .

Les hypothèses ci-dessus entrainent d'après R.F. Bass et E. Pardoux [1], l'existence d'une unique probabilité P^X sur $(\Omega_1, \mathcal{F}_1)$, solution du problème de martingale associé à l'équation (2 1) avec condition initiale Π_0. Il existe alors un processus $\{U_t\,,\, t \geq 0\}$ P^X-Wiener standard à valeurs dans $\mathbb{R}^n$ tel que $\{X_t\,,\, t \geq 0\}$ est donné par (2.1). On définit sur Ω la probabilité suivante : $P = P^X \otimes \bar{P}$ avec $\bar{P}$ la mesure du Wiener sur $(\Omega_2, \mathcal{F}_2)$. Notre espace de probabilité sera alors $(\Omega\,, \mathcal{F}\,, \mathcal{F}_t\,, P)$. On remarque que le processus d'observation $\{Y_t\,,\, t \geq 0\}$ défini par (2.2) dépend de ε , mais pour simplifier les notations nous omettons de le noter en index .

Réécrivons le problème précédent sous une forme plus usuelle :

$$\left\{ \begin{array}{rcl} X_t &=& X_0 + \int_0^t b(X_s)ds + \int_0^t f(X_s)dV_s + \int_0^t g(X_s)dW_s, \\ Y_t &=& \int_0^t h(X_s)ds + \varepsilon W_t, \end{array} \right. \tag{2.3}$$

où $\{(V_t, W_t)^*\,,\, t \geq 0\}$ est un processus de Wiener standard à valeurs dans $\mathbb{R}^{2n}$ défini par

$$\begin{array}{rcl} V_t &=& (I_n - \Upsilon^*\Upsilon)^{-1/2}(U_t - \Upsilon^* W_t) \\ W_t &=& \Upsilon U_t + \bar{\Upsilon}\bar{U}_t \end{array}$$

et

$$\begin{array}{rcl} f(X_t) &=& \sigma(X_t)(I_n - \Upsilon^*\Upsilon)^{1/2}, \\ g(X_t) &=& \sigma(X_t)\Upsilon^*. \\ F_i F_i^* + G_i G_i^* &=& \Sigma_i \Sigma_i^*,\ \forall i = 1, \cdots, l. \end{array}$$

D'après les hypothèses $(H4)$ et $(H5)$, la matrice F_i est inversible, $\forall i = 1, \cdots, l$.

Pour $i = 1, \cdots, l$ considérons le filtre de Kalman-Bucy (FK_i), le filtre optimal du problème de filtrage linéaire (i):

$$\begin{cases} X_t &= \eta + \int_0^t (B_i X_s + b_i)ds + \int_0^t F_i dV_s + \int_0^t G_i dW_s, \\ Y_t &= \int_0^t (H_i X_s + h_i)ds + \varepsilon W_t, \end{cases}$$

où η est un vecteur aléatoire suivant une loi normale de moyenne et de variance égales à celles de la condition initiale de l'équation (2.1).

Soit $\hat{X}_t^i$, l'espérance conditionnelle de X_t et $R_t^{i,\varepsilon}$ la matrice de covariance conditionnelle, pour tout $t \geq 0$, $R_t^{i,\varepsilon}$ est solution de l'équation de Riccati:

$$\dot{R}_t^{i,\varepsilon} = \Sigma_i \Sigma_i^* + B_i R_t^{i,\varepsilon} + R_t^{i,\varepsilon} B_i^* - (K_t^i)(K_t^i)^* \tag{2.4}$$

avec $K_t^i = \frac{1}{\varepsilon} R_t^{i,\varepsilon} H_i^* + G_i$. L'équation (2.4) se réécrit:

$$\dot{R}_t^{i,\varepsilon} = F_i F_i^* + (B_i - \frac{1}{\varepsilon} G_i H_i) R_t^{i,\varepsilon} + R_t^{i,\varepsilon} (B_i - \frac{1}{\varepsilon} G_i H_i)^* - \frac{1}{\varepsilon^2} R_t^{i,\varepsilon} H_i^* H_i R_t^{i,\varepsilon}. \tag{2.5}$$

D'après W. Murray Wonham (cf.[15] [th12.2]), les hypothèses $(H4)$ et $(H5)$ assurent l'existence et l'unicité d'une solution stationnaire $R_\infty^{i,\varepsilon} > 0$. On a :

$$F_i F_i^* + (B_i - \frac{1}{\varepsilon} G_i H_i) R_\infty^{i,\varepsilon} + R_\infty^{i,\varepsilon} (B_i - \frac{1}{\varepsilon} G_i H_i)^* - \frac{1}{\varepsilon^2} R_\infty^{i,\varepsilon} H_i^* H_i R_\infty^{i,\varepsilon} = 0 \tag{2.6}$$

et une formule explicite pour l'équation (2.5) :

$$R_t^{i,\varepsilon} = R_\infty^{i,\varepsilon} + e^{\frac{1}{\varepsilon} A_i t} \left[(R_0^{i,\varepsilon} - R_\infty^{i,\varepsilon})^{-1} + \int_0^t e^{\frac{1}{\varepsilon} A_i s} (H_i^* H_i) e^{\frac{1}{\varepsilon} A_i^* s} ds \right]^{-1} e^{\frac{1}{\varepsilon} A_i^* t},$$

avec $A_i = \varepsilon B_i - K_\infty^i H_i$ et $K_\infty^i = \frac{1}{\varepsilon} R_\infty^{i,\varepsilon} H_i^* + G_i$. Introduisons la définition suivante:

Définition 2.1 *Une matrice M appartenant à $\mathbb{R}^n \otimes \mathbb{R}^n$ est dite stable, si toutes les parties réelles de ses valeurs propres sont strictement négatives.*

Vu que le système (i) considéré est observable et commandable, d'après [15] [th12.2], la matrice $A_i = \varepsilon B_i - K_\infty^i H_i$ est stable.

L'intégrale $\int_0^t e^{\frac{1}{\varepsilon} A_i s} (H_i^* H_i) e^{\frac{1}{\varepsilon} A_i^* s} ds$ est donc bornée sur $t \in [0, +\infty[$ et on obtient la vitesse de convergence vers la solution stationnaire :

$$|R_t^{i,\varepsilon} - R_\infty^{i,\varepsilon}| \leq k e^{-k\mu_i t/\varepsilon}, \ k > 0. \tag{2.7}$$

avec $\mu_i > 0$ tel que pour un certain $\varepsilon_0 > 0$,

$$\sup_{0\leq\varepsilon\leq\varepsilon_0} \{ \textit{parties réelles des valeurs propres de la matrice } A_i\} \leq -\mu_i < 0,$$

Montrons que la matrice $R^{i,\varepsilon}_\infty$ est d'ordre ε. D'après l'équation (2.6), étant donné que la matrice H_i est inversible, il existe k_1, k_2 et $k_3 > 0$ tels que la norme de $R^{i,\varepsilon}_\infty$ vérifie l'inégalité

$$k_1 \, |R^{i,\varepsilon}_\infty|^2 \leq k_2\varepsilon|R^{i,\varepsilon}_\infty| + k_3\varepsilon^2.$$

D'où, $|R^{i,\varepsilon}_\infty| < k\varepsilon$ pour ε petit et $k > 0$ et il existe une sous-suite de $\frac{1}{\varepsilon}R^{i,\varepsilon}_\infty$ qui converge lorsque $\varepsilon \to 0$. Or, toute sous-suite convergente de $\frac{1}{\varepsilon}R^{i,\varepsilon}_\infty$ a sa matrice limite R^i_∞ solution de l'équation de Riccati:

$$F_iF_i^* - G_iH_iR^i_\infty - R^i_\infty H_i^*G_i^* - R^i_\infty H_i^*H_iR^i_\infty = 0. \tag{2.8}$$

Les hypothèses $(H3)$ et $(H4)$ impliqucnt que l'équation (2.8) admet une unique solution définie positive et la matrice $-G_iH_i - R^i_\infty H_i^*H_i$ est stable. On en déduit que

$$R^{i,\varepsilon}_\infty = \varepsilon R^i_\infty + R^i_\varepsilon, \tag{2.9}$$

avec $R^i_\varepsilon \sim o(\varepsilon)$. L'équation (2.8) se réécrit:

$$\Sigma_i\Sigma_i^* = (R^i_\infty H_i^* + G_i)(R^i_\infty H_i^* + G_i)^* \tag{2.10}$$

et nous avons :

$$H_i\Sigma_i\Sigma_i^*H_i^* = H_i(R^i_\infty H_i^* + G_i)(R^i_\infty H_i^* + G_i)^*H_i^*. \tag{2.11}$$

Remarque 2.2 L'équation (2.11) admet une unique solution symétrique définie positive $P_i = H_iR^i_\infty H_i^*$, de plus la matrice $-H_iG_i - H_iR^i_\infty H_i^*$ est stable.

Lorsque $\varepsilon \to 0$, les filtres de Kalman-Bucy (FK_i) considérés étant à "mémoire courte", le choix de la loi initiale est sans importance. Pour simplifier, nous prennons la loi $N(\mathbf{E}(X_0), R^{i,\varepsilon}_\infty)$. Le filtre de Kalman-Bucy associé à cette loi initiale est donné par:

$$\begin{cases} d\hat{X}^i_t &= (B_i\hat{X}^i_t + b_i)dt + \frac{1}{\varepsilon}K^i_\infty(dY_t - (H_i\hat{X}^i_t + h_i)dt), \\ \hat{X}^i_0 &= \mathbf{E}(X_0). \end{cases} \tag{2.12}$$

Soit $(e_1, e_2, .., e_n)$ la base canonique de l'espace vectoriel $\mathbb{R}^n$, munissons l'espace affine euclidien $\mathcal{R}^n$ du repère orthonormé $(0, e_1, e_2, .., e_n)$ et nous convenons de noter indifféremment par la lettre x, le point affine x et le vecteur $0x$ de l'espace vectoriel sous-jacent. Ainsi, si x est un point de l'espace affine $\mathcal{R}^n$ et si u est un vecteur de $\mathbb{R}^n$, $\langle x\,,\,u\rangle$ désigne le produit scalaire des vecteurs $0x$ et u. Par la suite, nous désignerons par $\mathbb{R}^n$ l'espace affine et l'espace vectoriel. Si $u \in \mathbb{R}^n$, $u(i)$, $1 \leq i \leq n$ représentera la i-ième composante du vecteur u et $|u|$ sa norme euclidienne. Si $M \in \mathbb{R}^n \otimes \mathbb{R}^d$, M_{ij} , $1 \leq i \leq n, 1 \leq j \leq d$, représentera l'élément de la i-ième ligne et de la j-ième colonne de la matrice M et nous prendrons pour norme $|M| = \sqrt{\text{Trace } MM^*}$ où M^* désigne la transposée de la matrice M.

On utilisera également les notations suivantes: $\|B\| = \sup(|B_-|, |B_+|)$, $\|H\| = \sup(|H_-|, |H_+|)$ et $\|\Sigma\| = \sup(|\Sigma_-|, |\Sigma_+|)$, et la définition

Définition 2.3 *Une variable aléatoire X sera dite d'ordre $\mathcal{O}(\varepsilon^q)$, $q \geq 0$, s'il existe $\varepsilon_0 > 0$, $c = c(\varepsilon_0) > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$:*

$$\varepsilon^{-2q}\mathbf{E}(|X|^2) \leq \mathrm{c}.$$

Pour simplifier, nous présentons la procédure dans le cas $l = 2$ et nous étendrons nos résultats au cas $l \geq 2$ dans le paragraphe 8. De plus, nous supposons que :

$$b_1 = b_2 = h_1 = h_2 = 0.$$

On remarque que cette hypothèse est non restrictive vu que les fonctions b et h sont supposées continues.

Nous supposons donc, sans perte de généralité que les deux demi-espaces Θ_1 et Θ_2 sont définis de la façon suivante:

Soit u un vecteur normé de $\mathbb{R}^n$, $|u| = 1$, Δ_u l'hyperplan séparateur des deux demi-espaces Θ_1 et Θ_2 :

$$\begin{aligned}
\Delta_u &= \{x \in \mathbb{R}^n\,,\,\langle x\,,\,u\rangle = 0\} \\
\Theta_1 &= \{x \in \mathbb{R}^n\,,\,\langle x\,,\,u\rangle < 0\} \\
\Theta_2 &= \{x \in \mathbb{R}^n\,,\,\langle x\,,\,u\rangle \geq 0\}.
\end{aligned}$$

Nous remplaçons les indices 1 et 2 respectivement par $-$ et $+$.

Introduisons la matrice de projection de $\mathbb{R}^n$ sur Δ_u, notée Q:

$$\begin{aligned} Q &= I_n - uu^*, \\ Qu &= 0. \end{aligned}$$

Les vecteurs x, $b(x)$, $h(x)$ se réécrivent de la façon suivante :

$$x = \langle x\,,\,u\rangle u + Qx.$$

$$b(x) = \begin{cases} \langle x\,,\,u\rangle B_- u + B_- Qx & \text{si } \langle x\,,\,u\rangle \leq 0, \\ \langle x\,,\,u\rangle B_+ u + B_+ Qx & \text{si } \langle x\,,\,u\rangle \geq 0. \end{cases}$$

$$h(x) = \begin{cases} \langle x\,,\,u\rangle H_- u + H_- Qx & \text{si } \langle x\,,\,u\rangle \leq 0, \\ \langle x\,,\,u\rangle H_+ u + H_+ Qx & \text{si } \langle x\,,\,u\rangle \geq 0. \end{cases}$$

La continuité des applications b et h entraine:

$$B_- Q = B_+ Q, \tag{2.13}$$

$$H_- Q = H_+ Q. \tag{2.14}$$

La restriction de h à Δ_u étant une bijection de Δ_u sur $h(\Delta_u)$, nous pouvons supposer sans restreindre la généralité que :

$(H6)$ $\quad h(\Delta_u) = \Delta_u.$

Lemmes préliminaires

Lemme 2.4 *Soit Λ une matrice (n,n) symétrique définie positive, $U_1, U_2, ..., U_p$, p variables aléatoires i.i.d. à valeurs dans $\mathbb{R}^n$ de loi commune $\mathcal{N}(0,\Lambda)$), alors pour tout $c > 0$ on a :*

$$P\left(\left\{\max_{1\leq k\leq p} |U_k| \geq c\right\}\right) \leq p\,n \exp\{-c^2/2n \text{ Trace } (\Lambda)\}.$$

Preuve

$$P\left(\left\{\max_{k=1,\cdots,p}|U_k| \geq c\right\}\right) = 1 - P\left(\{|U_k| < c, 1 \leq k \leq p\}\right)$$

$$P\left(\{|U_k| < c, 1 \leq k \leq p\}\right) = \prod_{i=1}^{p} P(|U_k| < c) = (1 - P(|U_1| \geq c))^p$$

avec

$$P(|U_1| \geq c) \leq \sum_{i=1}^{n} P(|U_1(i)| \geq c/\sqrt{n}) \leq n \exp\{-c^2/2n \text{ Trace } (\Lambda)\}.$$

D'autre part, on remarque que pour $x \geq 0$ et $\alpha \geq 1$,

$$1 - (1 - x \wedge 1)^\alpha \leq \alpha x$$

d'où le résultat. □

Lemme 2.5 *Soit $(\Omega, \mathcal{F}, \mathcal{F}_t, P, W_t)$ un processus de Wiener standard à valeurs dans $\mathbb{R}^n$. On considère l'équation différentielle stochastique suivante :*

$$\begin{cases} dX_t &= BX_t dt + \Sigma dW_t, \\ X_0 &= \eta, \end{cases} \tag{2.15}$$

où B et Σ sont deux matrices (n,n) et η est un vecteur aleatoire $\mathcal{F}_0$-mesurable à valeurs dans $\mathbb{R}^n$, indépendant du processus $\{W_t, t \geq 0\}$. On suppose qu'il existe $c_0 > 0$ tel que

$$\mathbf{E}(\exp c_0|\eta|^2) < +\infty, \tag{2.16}$$

Alors l'équation (2.15) admet une solution forte unique $\{X_t, t \geq 0\}$ et pour tout $T > 0$, il existe $c = c(T)$ tel que :

$$\mathbf{E}\left(\exp\left\{c \sup_{[0,T]}|X_t|^2\right\}\right) < +\infty.$$

Remarque 2.6 La condition (2.16) est satisfaite lorsque η est un vecteur aléatoire gaussien.

Preuve La condition de Lipschitz étant vérifiée par les coefficients de l'équation, l'existence et l'unicité en découlent immédiatement. La solution s'écrit :

$$X_t = e^{Bt}(\eta + \int_0^t e^{-Bs}\Sigma dW_s).$$

On pose $M_t = \int_0^t e^{-Bs}\Sigma dW_s$.

$$|X_t|^2 \leq |e^{Bt}|^2\,(2|\eta|^2 + 2|M_t|^2).$$

Soit $c \leq \frac{1}{2}c_0 \sup_{[0,T]} |e^{Bt}|^{-2}$, d'après la condition (2.16) et l'indépendance de η et $\{M_t\,,\, t \geq 0\}$, il suffit d'établir:

$$\mathbf{E}\left(\exp\left\{2c \sup_{[0,T]} |e^{Bt}|^2 |M_t|^2\right\}\right) < +\infty.$$

Il est clair que $\{M_t\,, t \geq 0\}$ est une martingale gaussienne centrée à valeurs dans $\mathbb{R}^n$. Il existe donc $c_1 > 0$ tel que

$$\mathbf{E}\left(\exp\{c_1|M_T|^2\}\right) < +\infty,$$

Le processus $\{\exp(\frac{c_1}{2}|M_t|^2)\,,\, t \geq 0\}$ étant une sous martingale, on a donc:

$$\begin{aligned}\mathbf{E}\left(\sup_{[0,T]} \exp\{c_1|M_t|^2\}\right) &= \mathbf{E}\left(\exp\left\{c_1 \sup_{[0,T]} |M_t|^2\right\}\right) \\ &\leq 4\mathbf{E}\left(\exp\{c_1|M_T|^2\}\right) < +\infty.\end{aligned}$$

On prend $c \leq \frac{1}{2}\sup_{[0,T]} |e^{Bt}|^{-2} \min(c_0, c_1)$ et on obtient le résultat demandé. □

Lemme 2.7 *Soit $(\Omega, \mathcal{F}, \mathcal{F}_t, P, W_t)$ un processus de Wiener standard à valeurs dans $\mathbb{R}^n$. Si le processus $\{X_t\,, t \geq 0\}$ vérifie l'équation différentielle sochastique*

$$\left\{\begin{array}{rcl} dX_t &=& b(X_t)dt + \sigma_t dW_t, \\ X_0 &=& \eta, \end{array}\right. \tag{2.17}$$

et que les hypothèses suivantes sont satisfaites:

i_1) $\mathbf{E}(\exp c_0|\eta|^2) < +\infty$ *pour un certain* $c_0 > 0$,

i_2) b est une application mesurable de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et il existe $k_1 > 0$ tel que:

$$|b(x)| \leq k_1(1 + |x|);$$

i_3) le processus $\{\sigma_t\,;\, t \geq 0\}$ est un processus $\mathcal{F}_t$-adapté à valeurs dans $\mathbb{R}^n \otimes \mathbb{R}^n$ et il existe $k_2 > 0$ tel que $|\sigma_t| \leq k_2$ p.s., t p.p.,

alors pour tout $T > 0$, il existe $c = c(T) > 0$ tel que

$$\mathbf{E}\left(\exp\left\{c \sup_{[0,T]} |X_s|^2\right\}\right) < +\infty.$$

Preuve En adaptant la démonstration donnée par R.S. Liptzer et A.N. Shyriaev [9][th 4.7] au cas $n > 1$, on montre sans difficulté que sous les hypothèses i_1), i_2) et i_3), pour tout $T > 0$, il existe $c_1 = c_1(T) > 0$ tel que

$$\sup_{[0,T]} \mathbf{E}\left(\exp\{c_1|X_t|^2\}\right) < +\infty. \tag{2.18}$$

Par la formule d'itô, nous obtenons:

$$\begin{aligned}
&\exp\{c|X_t|^2\} \\
= \;& \exp\{c|X_0|^2\} + 2c\int_0^t \exp\{c|X_s|^2\} X_s^* b(X_s)\, ds \\
&+2c\int_0^t \exp\{c|X_s|^2\} X_s^* \sigma_s\, dW_s \\
&+c\int_0^t \exp\{c|X_s|^2\} \text{ Trace } (\sigma_s\sigma_s^* + 2cX_sX_s^*\sigma_s\sigma_s^*)\, ds.
\end{aligned}$$

On en déduit l'inégalité suivante:

$$\begin{aligned}
\mathbf{E}\left(\sup_{[0,T]} \exp\{c|X_s|^2\}\right) \;\leq\; & \mathbf{E}\left(\exp\{c|X_0|^2\}\right) \\
&+2ck_1\int_0^T \mathbf{E}\left(\left[\exp\{c|X_s|^2\}\right]\left[|X_s| + |X_s|^2\right]\right) ds \\
&+2c\mathbf{E}\left(\sup_{[0,T]} \int_0^t \exp\{c|X_s|^2\} X_s^* \sigma_s\, dW_s\right) \\
&+ck_2^2\int_0^T \mathbf{E}\left(\left(\exp\{c|X_s|^2\}\right)\left(1 + 2c|X_s|^2\right)\right) ds.
\end{aligned}$$

l'inégalité de Cauchy-Schwartz et la majoration (2.18) nous permet d'utiliser l'inégalité de Burkholder-Davis-Gundy pour traiter le troisième terme de droite de l'inégalité. Ainsi, si nous choisissons $c \leq \min(c_0, c_1/2)$, nous pouvons conclure

$$\mathbf{E}\left(\exp\left\{c \sup_{[0,T]} |X_s|^2\right\}\right) < +\infty.$$

□

Lemme 2.8 *Soit $(\Omega, \mathcal{F}, \mathcal{F}_t, P, W_t)$ un processus de Wiener standard à valeurs dans $\mathbb{R}^n$, $\{\varphi_t, t \geq 0\}$ un processus $\mathcal{F}_t$-mesurable à valeurs dans $\mathbb{R}^n$ tel que pour tout $T > 0$,*

$$\mathbf{E}\left(\exp\left\{c_0 \sup_{[0,T]} |\varphi_t|^2\right\}\right) < +\infty, \text{ pour un certain } c_0 > 0. \tag{2.19}$$

Si le processus $\{Z_t^\varepsilon, t \geq 0\}$ vérifie l'équation différentielle stochastique suivante:

$$\begin{cases} dZ_t^\varepsilon &= \frac{1}{\varepsilon}(\Lambda Z_t^\varepsilon + \varphi_t)\, dt + \Sigma dW_t, \\ Z_0^\varepsilon &= z_0, \end{cases}$$

avec $\varepsilon > 0$, $z_0 \in \mathbb{R}^n$, Σ une matrice (n,n) et Λ une matrice (n,n) stable, alors pour tout $T > 0$, il existe $c > 0$ tel que :

$$\limsup_{\varepsilon \to 0} \mathbf{E}\left(\exp\left\{c \sup_{[0,T]} |Z_t^\varepsilon|^2\right\}\right) < +\infty.$$

Preuve

$$Z_t = e^{\frac{1}{\varepsilon}\Lambda t} z_0 + \frac{1}{\varepsilon}\int_0^t e^{\frac{1}{\varepsilon}\Lambda(t-s)} \varphi_s\, ds + \int_0^t e^{\frac{1}{\varepsilon}\Lambda(t-s)} \Sigma\, dW_s.$$

La matrice Λ étant stable, il existe $\lambda > 0$, $\delta > 0$ tels que:

$$|e^{\frac{1}{\varepsilon}\Lambda t}| \leq \delta e^{-\frac{\lambda}{\varepsilon}t}, \qquad \forall t \geq 0. \tag{2.20}$$

On obtient:

$$\sup_{[0,T]} |Z_t| \leq \delta |z_0| + \frac{\delta}{\lambda} \sup_{[0,T]} |\varphi_t| + \sup_{[0,T]} |K_t|,$$

avec

$$K_t = \int_0^t e^{\frac{1}{\varepsilon}\Lambda(t-s)} \Sigma\, dW_s.$$

D'après l'inégalité $|X+Y|^2 \leq 2|X|^2 + 2|Y|^2$ et celle de Cauchy-Schwartz, il suffit de considérer les trois termes séparément.

Le premier est déterministe et indépendant de ε. La condition (2.19) assure un moment exponentiel borné indépendament de ε pour le second. Considérons le troisième terme. En dimension un, on peut conclure rapidement à l'aide du théorème de comparaison des diffusions [7] et du lemme 2.5. En dimension supérieure à un, un petit travail s'impose.

Utilisons le fait que la matrice Λ est stable. L'équation

$$\Lambda^* \Pi + \Pi \Lambda = -I_n, \tag{2.21}$$

admet alors une unique solution symétrique définie positive donnée par

$$\Pi = \int_{-\infty}^{+\infty} e^{\Lambda^* t} e^{\Lambda t} dt.$$

Vu que $\forall c > 0, \exists c_1 > 0$ tel que $c_1 I_n - c\Pi < 0$, il suffit de montrer qu'il existe $c > 0$ tel que

$$\limsup_{\varepsilon \to 0} \mathbf{E}\left(\sup_{[0,T]} \exp\{cK_t^* \Pi K_t\}\right) < +\infty.$$

Par la formule d'Itô et (2.21), on obtient:

$$\begin{aligned}
\exp\{cK_t^* \Pi K_t\} &= 1 - \frac{c}{\varepsilon}\int_0^t \exp\{cK_s^* \Pi K_s\} K_s^* K_s \, ds \\
&\quad + c\int_0^t \exp\{cK_s^* \Pi K_s\} \left[|\Pi^{\frac{1}{2}} \Sigma|^2 + 2c|K_s^* \Pi \Sigma|^2 \right] ds \\
&\quad + 2c\int_0^t \exp\{cK_s^* \Pi K_s\} K_s^* \Pi \Sigma dW_s.
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}\left(\sup_{[0,T]} \exp\{cK_t^* \Pi K_t\}\right) &\leq 1 + kT \sup_{[0,T]} \mathbf{E}\left(\exp\{cK_t^* \Pi K_t\}\right) \\
&\quad + 2kT \sup_{[0,T]} \left(\mathbf{E}\left(\exp\{2cK_t^* \Pi K_t\}\right)\right)^{\frac{1}{2}} \sup_{[0,T]} \left(\mathbf{E}|K_t|^4\right)^{\frac{1}{2}} \\
&\quad + 2c\mathbf{E}\left(\sup_{[0,T]} \int_0^t \exp\{cK_s^* \Pi K_s\} K_s^* \Pi \Sigma dW_s\right). \quad (2.22)
\end{aligned}$$

Ainsi, s'il existe $c > 0$, $k > 0$ tels que $\forall t \in [0,T]$,

$$\limsup_{\varepsilon \to 0} \mathbf{E}(\exp cK_t^* K_t) < k, \tag{2.23}$$

on conclut en utilisant l'inégalité de Burkholder-Davis-Gundy. Pour compléter la preuve du lemme, nous devons donc établir la majoration (2.23).

On remarque que $\{K_t, t \geq 0\}$ est un processus gaussien centré. Soit $V_t^\varepsilon \triangleq \mathbf{E}(K_t K_t^*)$. La matrice V_t^ε est solution de l'équation matricielle:

$$\dot{V}_t^\varepsilon = \frac{1}{\varepsilon}\Lambda V_t^\varepsilon + \frac{1}{\varepsilon} V_t^\varepsilon \Lambda^* + \Sigma\Sigma^* , \quad V_0^\varepsilon = 0.$$

$\forall 0 \leq t \leq T$, V_t^ε est une matrice symétrique semi-définie positive de rang d avec $0 \leq d \leq n$. Elle est donnée par

$$V_t^\varepsilon = e^{\frac{1}{\varepsilon}\Lambda t}\left(\int_0^t e^{-\frac{1}{\varepsilon}\Lambda s}\Sigma\Sigma^* e^{-\frac{1}{\varepsilon}\Lambda^* s}\, ds\right) e^{\frac{1}{\varepsilon}\Lambda^* t}.$$

La matrice Λ étant stable, V_t^ε est d'ordre ε. Plus précisément, Il existe $\varepsilon_0 > 0$, $k_1 > 0$ tel que $\forall \varepsilon \in (0, \varepsilon_0]$, $\forall t \in [0,T]$,

$$|V_t^\varepsilon| < \varepsilon k_1. \tag{2.24}$$

Raisonnons à t fixé: la variable aléatoire K_t suit une loi normale $\mathcal{N}(0, V_t^\varepsilon)$. Il existe une matrice unitaire U tel que

$$UV_t^\varepsilon U^* = \begin{pmatrix} \bar{V}_t^\varepsilon & 0 \\ 0 & 0 \end{pmatrix}$$

avec $\bar{V}_t^\varepsilon$ une matrice diagonale (d,d) de rang d et

$$UK_t = \begin{pmatrix} \bar{K}_t \\ 0 \end{pmatrix}$$

avec $\bar{K}_t$ une variable aléatoire gaussienne centrée à valeurs dans $\mathbb{R}^d$, i.e. $\bar{K}_t \sim \mathcal{N}(0, \bar{V}_t^\varepsilon)$. Vu que $|K_t| = |\bar{K}_t|$, on s'intéresse à $\mathbf{E}(\exp c|\bar{K}_t|^2)$.

Notons $v_t^1, \cdots, v_t^d$, les d valeurs propres strictement positives de la matrice V_t^ε. D'après (2.24), il existe $v > 0$ tel que:

$$\max(v_t^1, \cdots, v_t^d) \leq \varepsilon v.$$

On obtient:

$$\begin{aligned}\mathbf{E}(\exp\{c|\bar{K}_i|^2\}) &= \frac{1}{(2\pi)^{\frac{d}{2}}(\prod_{i=1}^{d} v_i^i)^{\frac{1}{2}}} \prod_{i=1}^{d} \int_{-\infty}^{+\infty} \exp\{(c - \frac{1}{2v_i^i})x_i^2\}\, dx_i \\ &\leq \left(\frac{1}{-2cv+1}\right)^{\frac{d}{2}}, \qquad \forall c < \frac{1}{2v}, \ \forall 0 < \varepsilon \ll 1.\end{aligned}$$

La constante c et le majorant ne dépendant ni de t, ni de ε, l'inégalité (2.23) est donc vérifiée.

□

3 Détection des passages de $\{X_t\}$ par Δ_u

Soit Δ_u, l'hyperplan séparateur des deux demi-espaces Θ_- et Θ_+,

$$\Delta_u = \{x \in \mathbb{R}^n ; \langle x , u \rangle = 0\}.$$

Nous devons premièrement déterminer sur l'intervalle de temps $[0, T]$, $T > 0$ fixé, des intervalles durant lesquels une trajectoire $t \to X_t$ ne passe pas par Δ_u avec une probabilité proche de 1. Nous proposons un test sur la sortie du filtre de Kalman (FK_+). Ce test s'applique indifféremment à la sortie du filtre de kalman (FK_-). Notre première étape consiste à montrer que pour de faibles valeurs de ε, la différence $h(X_t) - H_+\hat{X}_t^+$ est petite avec une probabilité proche de 1.

Proposition 3.1 *Soit* $0 < a < b$, $\forall \theta > 0$, $\forall \alpha \in [0, \frac{1}{2}[$, $\forall \gamma \in]2\alpha, 1[$, *il existe* $k > 0$, $\varepsilon_0 > 0$ *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\sup_{[a,b]} |h(X_t) - H_+\hat{X}_t^+| > \theta\varepsilon^\alpha\right\}\right) \leq \exp\left\{-\frac{k}{\varepsilon^{\gamma-2\alpha}}\right\}.$$

Preuve L'application h n'étant pas de classe C^2, on ne peut pas appliquer la formule d'Itô à $h(X_t)$.Cependant à chaque composante de $h(X_t)$ on peut associer une application de $\mathbb{R}$ dans $\mathbb{R}$ linéaire par morceaux.

Pour $1 \leq i \leq n$:

$$[h(x)](i) = \begin{cases} [H_-u](i)\langle x , u\rangle + [H_-Qx](i) & \text{si } \langle x , u\rangle \leq 0, \\ [H_+u](i)\langle x , u\rangle + [H_+Qx](i) & \text{si } \langle x , u\rangle > 0. \end{cases}$$

D'après (2.14) on obtient:

$$[h(x)](i) = h^i(\langle x , u\rangle) + [H_+Qx](i),$$

avec

$$h^i : \mathbb{R} \to \mathbb{R}$$
$$r \mapsto h^i(r) = \begin{cases} [H_-u](i)r & \text{si } r \leq 0, \\ [H_+u](i)r & \text{si } r \geq 0. \end{cases}$$

On pose $X_t^1 = \langle X_t , u\rangle$. D'après la formule d'Itô-Tanaka [8], $h^i(X_t^1)$ s'écrit:

$$h^i(X_t^1) = h^i(X_0^1) + \int_0^t h_-^{i\prime}(X_s^1)dX_s^1 + \frac{1}{2}[(H_+ - H_-)u](i)L_t^0, \qquad (3.1)$$

où

- $h_-^{i\prime}$ est la dérivée à gauche de h,

$$h_-^{i\prime}(r) = [H_-u](i)1_{\mathbb{R}^-}(r) + [H_+u](i)1_{\mathbb{R}^{+*}}(r).$$

- $\{L_t^0, t \geq 0\}$ est le processus temps local en 0 de la semi-martingale réelle $\{X_t^1, t \geq 0\}$.

On pose $Z_t = h(X_t) - H_+\hat{X}_t^+$. On obtient à partir de (2.12) et (3.1) l'équation suivante :

$$dZ_t = \frac{1}{\epsilon}\Lambda Z_t dt + \varphi_t dt + \psi_t^1 dV_t + \psi_t^2 dW_t + \frac{1}{2}(H_+ - H_-)uL_t^0. \tag{3.2}$$

avec

- $\Lambda = -H_+K_\infty^+$, matrice stable d'après la remarque (2.2).
- $\varphi_t = \langle b(X_t), u\rangle h'_-(X_t^1) - H_+B_+\hat{X}_t^+ + H_+Qb(X_t)$.
- $\psi_t^1 = H_+Qf(X_t) + h'_-(X_t^1)u^*f(X_t)$.
- $\psi_t^2 = H_+(Qg(X_t) - K_\infty^+) + h'_-(X_t^1)u^*g(X_t)$
- $h'_-(X_t^1)^* = (h_-^{1\prime}(X_t^1), h_-^{2\prime}(X_t^1), \ldots, h_-^{n\prime}(X_t^1))$

On décompose Z_t de la manière suivante :

$$Z_t = Z_t^{(1)} + Z_t^{(2)} + Z_t^{(3)} + Z_t^{(4)},$$

avec

- $Z_t^{(1)} = e^{\frac{1}{\epsilon}\Lambda t}Z_0$
- $Z_t^{(2)} = \int_0^t e^{\frac{1}{\epsilon}\Lambda(t-s)}\varphi_s ds$
- $Z_t^{(3)} = \int_0^t e^{\frac{1}{\epsilon}\Lambda(t-s)}(\psi_s^1 dVs + \psi_s^2 dWs)$
- $Z_t^{(4)} = \frac{1}{2}(H_+ - H_-)u\int_0^t e^{\frac{1}{\epsilon}\Lambda(t-s)}dL_s^0$

La matrice Λ étant stable, nous rappelons qu'il existe $\lambda > 0$, $\delta > 0$, tels que: $\forall t \geq 0$,

$$|e^{\frac{1}{\epsilon}\Lambda t}| \leq \delta e^{-\frac{\lambda}{\epsilon}t} \tag{3.3}$$

Considérons les quatre termes séparemment:

1) $Z_t^{(1)} = e^{\frac{1}{\epsilon}\Lambda t} Z_0$. En utilisant (3.3), on obtient:

$$|Z_t^{(1)}| \leq \delta e^{-\frac{\lambda}{\epsilon}a} |Z_0|, \ \forall\, t \in [a,b].$$

D'après l'inégalité de Bienaymé-Tchebichev :

$$\begin{aligned} P\left(\left\{\sup_{[a,b]} |Z_t^{(1)}| > \frac{\theta\varepsilon^\alpha}{4}\right\}\right) &\leq P\left(\{|Z_0| > \frac{\theta\varepsilon^\alpha e^{\frac{\lambda}{\epsilon}a}}{4}\}\right) \\ &\leq \frac{4}{\theta\varepsilon^\alpha} e^{-\frac{\lambda}{\epsilon}a} \mathbf{E}|Z_0|, \end{aligned}$$

et la majoration suivante,

$$\mathbf{E}|Z_0| \leq 2\|H\|\,\mathbf{E}|X_0| < +\infty,$$

on conclut l'existence de $k_1 > 0$, $\varepsilon_1 > 0$ tels que :

$$P\left(\left\{\sup_{[a,b]} |Z_t^{(1)}| > \frac{\theta\varepsilon^\alpha}{4}\right\}\right) \leq e^{-k_1/\epsilon},\ 0 < \varepsilon \leq \varepsilon_1. \tag{3.4}$$

2) $Z_t^{(2)} = \int_0^t e^{\frac{1}{\epsilon}\Lambda(t-s)} \varphi_s ds$. On remarque que :

$$|\varphi_t| \leq \|B\|\,(\,\|H\| + |H_+Q|\,)\,|X_t| + |H_+B_+||\hat{X}_t^+|$$

Les lemmes 2.7 et 2.8 impliquent que pour tout $c > 0$, $T \geq 0$, il existe $k_1 = k_1(T,c) > 0$ et $k_2 = k_2(T,c) > 0$ tels que

$$\mathbf{E}\left(\exp\left\{c \sup_{[0,T]} |X_t|\right\}\right) < k_1,$$

$$\limsup_{\epsilon \to 0} \mathbf{E}\left(\exp\left\{c \sup_{[0,T]} |\hat{X}_t^+|\right\}\right) < k_2.$$

On en déduit que :

$$\mathbf{E}\left(\exp\left\{c \sup_{[0,T]} |\varphi_t|\right\}\right) < \infty \text{ pour } c > 0. \tag{3.5}$$

D'après (3.5) et l'inégalité $|Z_t^{(2)}| \leq \frac{\delta\varepsilon}{\lambda} \sup_{[0,t]} |\varphi_s|$, il existe $k_2 > 0$, $\varepsilon_2 > 0$ tels que

$$P\left(\left\{\sup_{[a,b]} |Z_t^{(2)}| > \frac{\theta\varepsilon^\alpha}{4}\right\}\right) \leq \exp\left\{-\frac{k_2}{\varepsilon^{1-\alpha}}\right\}, \qquad 0 < \varepsilon \leq \varepsilon_2. \tag{3.6}$$

3) $Z_t^{(3)} = \int_0^t e^{\frac{1}{\varepsilon}\Lambda(t-s)}(\psi_s^1 dVs + \psi_s^2 dW_s)$.
On pose

$$\psi_t = \begin{pmatrix} \psi_t^1 \\ \psi_t^2 \end{pmatrix}.$$

On remarque qu'il existe $r > 0$ tel que $|\psi_t|^2 \leq r$. Le raisonnement qui suit est similaire à celui utilisé par W.H. Fleming et E. Pardoux dans la démonstration du lemme 3.1 [6]. L'idée est de construire une surmartingale positive $\{M(t, Z_t^{(3)})\,;\, 0 \leq t \leq b\}$:

$$M(t, Z_t^{(3)}) = e^{\rho t}\kappa(Z_t^{(3)}) + \bar{\rho}(e^{\rho b} - e^{\rho t})$$

où ρ et $\bar{\rho}$ sont à choisir de façon adéquate et l'application κ est définie par:

$$\begin{aligned} \kappa :\ & \mathbb{R}^n \to \mathbb{R} \\ & z \mapsto \kappa(z) = \exp\left\{\frac{1}{4rp\varepsilon}\, z^* \Pi z\right\}, \end{aligned}$$

où Π est une matrice symétrique définie positive, solution de l'équation

$$\Lambda^* \Pi + \Pi \Lambda = -I_n \tag{3.7}$$

et $p =$ Trace (Π^2). Soit $\mathcal{L}_t$, l'opérateur linéaire défini par

$$\mathcal{L}_t\kappa(z) = \frac{1}{\varepsilon} z^* \Lambda^* \kappa'(z) + \frac{1}{2} \text{ Trace } (\psi_t \psi_t^* \kappa''(z)).$$

En utilisant (3.7), on obtient

$$\begin{aligned} \mathcal{L}_t\kappa(z) &= \left[-\frac{1}{4\varepsilon^2 rp}|z|^2 + \frac{1}{4\varepsilon rp} \text{ Trace } (\Pi \psi_t \psi_t^*) \right. \\ &\qquad \left. + \frac{1}{8p^2 r^2 \varepsilon^2} \text{ Trace } (\Pi z^* z \Pi \psi_t \psi_t^*) \right] \kappa(z) \\ &\leq \frac{1}{\varepsilon}\left[-\frac{1}{8\varepsilon rp}|z|^2 + \frac{\text{Trace } (\Pi)}{4p} \right] \kappa(z). \end{aligned}$$

Vu que Trace $(\Pi) \leq \sqrt{n \text{ Trace } (\Pi^2)} = \sqrt{n\,p}$, on a:

$$\mathcal{L}_t\kappa(z) + \frac{1}{\varepsilon}\kappa(z) \leq \frac{1}{\varepsilon}\left[1 - \frac{1}{8\varepsilon rp}|z|^2 + \frac{\sqrt{n}}{4\sqrt{p}} \right] \kappa(z).$$

Donc, si $|z|^2 \geq 2\varepsilon r\,(\sqrt{n\,p} + 4p)$,

$$\mathcal{L}_t\kappa(z) + \frac{1}{\varepsilon}\kappa(z) \leq 0.$$

On en déduit la majoration suivante:

$$\mathcal{L}_t\kappa(z) + \frac{1}{\varepsilon}\kappa(z) \leq \frac{\left(4p + \sqrt{n\,p}\right)}{4\varepsilon\,p} \exp\left\{\frac{\left(n + 4\sqrt{n\,p}\right)}{2}\right\}.$$

On note $n_0 = \frac{\left(4p+\sqrt{n\,p}\right)}{4\,p} \exp\left\{\frac{\left(n+4\sqrt{n\,p}\right)}{2}\right\}$. Utilisons ce résultat pour déterminer ρ et $\bar{\rho}$.

$$\begin{aligned}\frac{\partial M}{\partial t} + \mathcal{L}_t M &= e^{\rho t}(\rho\kappa + \mathcal{L}_t\kappa - \bar{\rho}\rho) \\ &\leq e^{\rho t}\left(\left(\rho - \frac{1}{\varepsilon}\right)\kappa + \frac{n_0}{\varepsilon} - \bar{\rho}\rho\right).\end{aligned}$$

Ainsi, en choisissant $\rho = \frac{1}{\varepsilon^{\frac{1}{2}-\alpha}}$ et $\bar{\rho} = \frac{n_0}{\varepsilon^{\frac{1}{2}+\alpha}}$, pour $\varepsilon < 1$ et $0 \leq \alpha < \frac{1}{2}$, le processus $\{M(t, Z_t^{(3)});\ 0 \leq t \leq b\}$ est une surmartingale positive. Donc, pour tout $\theta > 0$, on a:

$$P\left(\sup_{[a,b]} M(t, Z_t^{(3)}) \geq \theta\right) \leq \frac{1 + \bar{\rho}(e^{\rho b} - 1)}{\theta}.$$

Etant donné que $|z|^2 \leq |\Pi^{-1}|\,|\Pi^{1/2}z|^2$ et $\forall 0 \leq t \leq b$, $\kappa(z) \leq M(t,z)$,

$$\begin{aligned}&P\left(\left\{\sup_{[a,b]} |Z_t^{(3)}| > \frac{\theta\varepsilon^{\alpha}}{4}\right\}\right) \\ \leq\ &P\left(\left\{\sup_{[a,b]} M(t, Z_t^{(3)}) > \exp\left\{\frac{\theta^2\varepsilon^{2\alpha}}{64\,|\Pi^{-1}|\,p\,r\,\varepsilon}\right\}\right\}\right) \\ \leq\ &\exp\left\{-\frac{\theta^2}{64\,|\Pi^{-1}|\,p\,r\,\varepsilon^{1-2\alpha}}\right\}\left[1 + \frac{n_0}{\varepsilon^{\frac{1}{2}+\alpha}}\left(\exp\left\{\frac{b}{\varepsilon^{\frac{1}{2}-\alpha}}\right\} - 1\right)\right] \\ \leq\ &\exp\left\{-\frac{k_3\theta^2}{\varepsilon^{1-2\alpha}}\right\}, \quad 0 < \varepsilon \leq \varepsilon_3, \qquad (3.3)\end{aligned}$$

pour un certain $k_3 > 0$ et $\varepsilon_3 > 0$.

4) $Z_t^{(4)} = \frac{1}{2}(H_+ - H_-)u \int_0^t e^{\frac{1}{\epsilon}\Lambda(t-s)} dL_s^0$. Rappelons que $\{L_t^0, t \geq 0\}$ est le processus temps local en 0 de la semi-martingale $\{X_t^1, t \geq 0\}$. Le processus $\{L_t^0, t \geq 0\}$ est positif et croissant, d'où l'inégalité de Jensen et (3.3) impliquent

$$|Z_t^{(4)}| \leq \frac{|H_+ - H_-|\delta}{2} \int_0^t e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0.$$

Montrons premièrement que $\forall c > 0$, $t \geq 0$,

$$\mathbf{E}\left(\exp\left\{cL_t^0\right\}\right) < +\infty. \tag{3.9}$$

D'après la formule d'Itô-Tanaka, on a :

$$|X_t^1| = |X_0^1| + \int_0^t \text{signe}(X_s^1) dX_s^1 + L_t^0. \tag{3.10}$$

On pose

$$\begin{aligned} K_t &= |X_0^1| + \int_0^t \text{signe}(X_s^1) dX_s^1, \\ \bar{K}_t &= \int_0^t \text{signe}(X_s^1) u^* f(X_s) dV_s + \int_0^t \text{signe}(X_s^1) u^* g(X_s) dW_s. \end{aligned}$$

$$K_t = |X_0^1| + \int_0^t \text{signe}(X_s^1) u^* b(X_s) ds + \bar{K}_t. \tag{3.11}$$

Le processus $\{L_t^0, t \geq 0\}$ est solution du problème de Skorohod associé à (3.10), (cf.[2]), i.e.:

$$L_t^0 = \sup_{[0,t]} K_s^-, \tag{3.12}$$

où $K_t^- = \sup(0, -K_t)$. Donc $L_b^0 \leq \sup_{[0,b]} |K_s|$. D'autre part, d'après (3.11),

$$\sup_{[0,b]} |K_s| \leq |X_0^1| + \|B\| b \sup_{[0,b]} |X_s| + \sup_{[0,b]} |\bar{K}_s|$$

L'inégalité de Cauchy-Schwartz nous permet de considérer séparément les moments exponentiels des "sup" de chaque terme. En utilisant l'hypothèse $(H1)$, le fait que les fonctions f et g sont bornées et le lemme 2.7, on obtient que $\forall c > 0$

$$\mathbf{E}\left(\exp\left\{c \sup_{[0,b]} |K_s|\right\}\right) < +\infty.$$

On en déduit (3.9).

Soit $\gamma \in]2\alpha, 1[$, utilisons la décomposition suivante :

$$\int_0^t e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 = \int_0^{t-\epsilon^\gamma} e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 + \int_{t-\epsilon^\gamma}^t e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0.$$

En choisissant $\beta = \dfrac{\theta}{4\delta|H_+ - H_-|}$, on a:

$$\begin{aligned} P\left(\left\{\sup_{[a,b]} |Z_t^{(4)}| > \frac{\theta\epsilon^\alpha}{4}\right\}\right) \leq\ & P\left(\left\{\sup_{[a,b]} \int_0^{t-\epsilon^\gamma} e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 > \beta\epsilon^\alpha\right\}\right) \\ & + P\left(\left\{\sup_{[a,b]} \int_{t-\epsilon^\gamma}^t e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 > \beta\epsilon^\alpha\right\}\right). \end{aligned}$$

Considérons le premier terme du membre de droite de l'inégalité :

$$\int_0^{t-\epsilon^\gamma} e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 \leq e^{-\lambda/\epsilon^{1-\gamma}} L_t^0.$$

$$\begin{aligned} P\left(\left\{\sup_{[a,b]} \int_0^{t-\epsilon^\gamma} e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 > \beta\epsilon^\alpha\right\}\right) &\leq P\left(\left\{L_b^0 > \beta\epsilon^\alpha e^{\lambda/\epsilon^{1-\gamma}}\right\}\right) \\ &\leq C\exp\left\{-\beta\epsilon^\alpha e^{\lambda/\epsilon^{1-\gamma}}\right\}, \end{aligned}$$

d'après (3.9) et l'inégalité de Markov. Donc il existe $\epsilon_0 > 0$, $c > 0$, tels que $\forall \epsilon \in (0, \epsilon_0]$,

$$P\left(\left\{\sup_{[a,b]} \int_0^{t-\epsilon^\gamma} e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 > \beta\epsilon^\alpha\right\}\right) \leq e^{-c/\epsilon}. \tag{3.13}$$

Intéressons nous au second terme : Soit $0 < a < b$, $m = [\frac{b-a}{\epsilon^\gamma}]$, on pose $t_k = a + k\epsilon^\gamma$, $k = 0, \ldots, m$:

$$P\left(\left\{\sup_{[a,b]} \int_{t-\epsilon^\gamma}^t e^{-\frac{\lambda}{\epsilon}(t-s)} dL_s^0 > \beta\epsilon^\alpha\right\}\right) \leq P\left(\left\{\sup_{[a,b]} (L_t^0 - L_{t-\epsilon^\gamma}^0) > \beta\epsilon^\alpha\right\}\right)$$

$$\leq P\left(\left\{\max_{0\leq k\leq m} \sup_{[t_k, t_{k+1}]} (L_t^0 - L_{t-\epsilon^\gamma}^0) > \beta\epsilon^\alpha\right\}\right)$$

$$
\begin{aligned}
&\leq 2P\left(\left\{\max_{0\leq k\leq m}(L^0_{t_k}-L^0_{t_{k-1}})>\beta\varepsilon^\alpha/2\right\}\right)\\
&\leq 2P\left(\left\{\max_{0\leq k\leq m}\sup_{[t_{k-1},t_k]}|K_t-K_{t_{k-1}}|>\beta\varepsilon^\alpha/2\right\}\right)\\
&\leq 2P\left(\left\{\max_{0\leq k\leq m}\sup_{[t_{k-1},t_k]}\left|\int_{t_{k-1}}^t \mathrm{signe}(X^1_s)u^*b(X_s)ds\right|>\beta\varepsilon^\alpha/4\right\}\right)\\
&\quad +2P\left(\left\{\max_{0\leq k\leq m}\sup_{[t_{k-1},t_k]}|\bar K_t-\bar K_{t_{k-1}}|>\beta\varepsilon^\alpha/4\right\}\right).
\end{aligned}
$$

En utilisant l'inégalité de Markov et le lemme 2.7, on majore le premier terme:

$$
\begin{aligned}
&P\left(\left\{\max_{0\leq k\leq m}\sup_{[t_{k-1},t_k]}\left|\int_{t_{k-1}}^t \mathrm{signe}(X^1_s)u^*b(X_s)ds\right|>\beta\varepsilon^\alpha/4\right\}\right)\\
&\leq P\left(\left\{\sup_{[a,b]}|X_s|>\frac{\beta^2}{4\|B\|\varepsilon^{\gamma-\alpha}}\right\}\right)\leq C\,e^{-c/\varepsilon^{\gamma-\alpha}}.
\end{aligned}
$$

Considérons le second terme:

$$
\begin{aligned}
&P\left(\left\{\max_{0\leq k\leq m}\sup_{[t_{k-1},t_k]}|\bar K_s-\bar K_{t_{k-1}}|>\beta\varepsilon^\alpha/4\right\}\right)\\
&\leq 1-P\left(\left\{\max_{0\leq k\leq m}\sup_{[t_{k-1},t_k]}|\bar K_s-\bar K_{t_{k-1}}|\leq\beta\varepsilon^\alpha/4\right\}\right).
\end{aligned}
$$

Pour $k=0,\ldots,m$, on pose

- $N_k=\sup_{[t_{k-1},t_k]}|\bar K_t-\bar K_{t_{k-1}}|$,
- $D_k=\{N_k\leq\beta\varepsilon^\alpha/4\}$,
- $\mathcal{H}_k=\sigma(D_i\,,\,i=0,k-1)$.

On remarque que $\mathcal{H}_k\subset\mathcal{F}_{t_k}$.

$$
\begin{aligned}
P\left(\left\{\max_{0\leq k\leq m}N_k\leq\beta\varepsilon^\alpha/4\right\}\right)&=P\left(\bigcap_{k=0}^m D_k\right)\\
&=E\left[\mathbf{1}_{D_0}\mathbf{1}_{D_1}\ldots\mathbf{1}_{D_{m-1}}\mathbf{E}(\mathbf{1}_{D_m}\,|\,\mathcal{H}_{m-1})\right].
\end{aligned}
$$

On s'intéresse à $\mathbf{E}(\mathbf{1}_{D_k} \mid \mathcal{H}_{k-1})$. Soit $H \in \mathcal{H}_{k-1}$, $\delta > 0$, on introduit la martingale exponentielle $\{M_t^{\delta,k-1}, t \geq t_{k-1}\}$:

$$M_t^{\delta,k-1} = \mathbf{1}_H \exp\left\{ \delta(\bar{K}_t - \bar{K}_{t_{k-1}}) - \frac{\delta^2}{2}\left(\langle \bar{K}\rangle_t - \langle \bar{K}\rangle_{t_{k-1}}\right)\right\}.$$

On note que pour $t_{k-1} \leq t \leq t_k$, $\langle \bar{K}\rangle_t - \langle \bar{K}\rangle_{t_{k-1}} \leq c\varepsilon^\gamma$ avec $c = \|F\|^2 + \|G\|^2$. D'après l'inégalité de Doob et en choisissant $\delta = \dfrac{\beta}{4c\varepsilon^{\gamma-\alpha}}$, on obtient:

$$\begin{aligned} & P\left(H \cap \left\{ \sup_{[t_{k-1},t_k]} (\bar{K}_t - \bar{K}_{t_{k-1}}) > \beta\varepsilon^\alpha/4 \right\}\right) \\ \leq\ & P\left(\left\{ \sup_{[t_{k-1},t_k]} M_t^{\delta,k-1} > \exp\left\{ \frac{\beta^2}{32c\varepsilon^{\gamma-2\alpha}} \right\}\right\}\right) \leq P(H)\exp\left\{ -\frac{\beta^2}{32c\varepsilon^{\gamma-2\alpha}} \right\}. \end{aligned}$$

De même, on montre qu'il existe $k > 0$, tel que

$$P\left(H \cap \left\{ \sup_{[t_{k-1},t_k]} (-\bar{K}_s + \bar{K}_{t_{k-1}}) > \beta\varepsilon^\alpha/4 \right\}\right) \leq P(H)e^{-k/\varepsilon^{\gamma-2\alpha}}.$$

On en déduit que $P(H \cap D_k^c) \leq P(H)e^{-k/\varepsilon^{\gamma-2\alpha}}$, $\forall H \in \mathcal{H}_{k-1}$. D'où ,

$$\mathbf{E}(\mathbf{1}_{D_k} \mid \mathcal{H}_k) \geq 1 - e^{-k/\varepsilon^{\gamma-2\alpha}},$$

$$\begin{aligned} P\left(\left\{ \max_{0\leq k\leq m} \sup_{[t_{k-1},t_k]} |K_s - K_{t_{k-1}}| \leq \beta\varepsilon^\alpha/4 \right\}\right) &= P(\bigcap_{k=0}^{m} D_k) \\ &\geq (1 - e^{-k/\varepsilon^{\gamma-2\alpha}})^{m+1}, \end{aligned}$$

et

$$\begin{aligned} P\left(\left\{ \sup_{[a,b]} \int_{t-\varepsilon^\gamma}^{t} e^{-\frac{\lambda}{\varepsilon}(t-s)} dL_s^0 > \beta \right\}\right) &\leq 2C\, e^{-c/\varepsilon^{2(\gamma-\alpha)}} + 2 - 2(1 - e^{-k/\varepsilon^{\gamma-2\alpha}})^{m+1} \\ &\leq 2C\, e^{-c/\varepsilon^{2(\gamma-\alpha)}} + 2(m+1)e^{-k/\varepsilon^{\gamma-2\alpha}} \quad (3.14) \end{aligned}$$

De (3.13) et (3.14) on conclut l'existence de $k_4 > 0$, $\varepsilon_4 > 0$ tels que pour $0 < \varepsilon \leq \varepsilon_4$,

$$P\left(\left\{ \sup_{[a,b]} |Z_t^{(4)}| > \frac{\theta\varepsilon^\alpha}{4} \right\}\right) \leq e^{-k_4/\varepsilon^{\gamma-2\alpha}}. \qquad (3.15)$$

Des inégalités (3.4), (3.6), (3.8) et (3.15), on obtient la proposition 3.1. □

Etant donnée que $h(\Delta_u^c) \cap h(\Delta_u) = \emptyset$, la proposition 3.1 nous permet de construire un test de détection des passages des trajectoires de $\{X_t\}$ par Δ_u basé sur $\{\hat{X}_t^+\}$.

Soit $0 < a < b$, on définit les événements suivants:

$$A_-(a,b) = \{\langle X_t\,,\, u\rangle < 0\,,\, a \leq t \leq b\},$$

$$A_+(a,b) = \{\langle X_t\,,\, u\rangle > 0\,,\, a \leq t \leq b\},$$

et pour $c > 0$ choisi,

$$C^\epsilon(a,b) = \{|\langle \hat{X}_t^+\,,\, u\rangle| > c\,,\, a \leq t \leq b\}.$$

Lorsqu'il n'y aura pas de confusion possible, ces événements seront notés: A_-, A_+, C^ϵ.

Remarque 3.2 D'après la proposition 3.1, nous pouvons remplacer la constante c utilisée dans l'événement test C^ϵ par $c\epsilon^\alpha$, avec $0 \leq \alpha < \frac{1}{2}$. Ceci permettrait d'agrandir les intervalles de monotonie détectés. Cependant, comme nous le verrons par la suite, sur les intervalles de monotonie, pour décider du signe de $\langle X_t\,,\, u\rangle$ sous les hypothèses du type (2), il est nécessaire que $\langle \hat{X}_t^+\,,\, u\rangle$ soit "suffisamment" éloigné de zéro durant un "certain" temps. Dans le cadre de ce chapitre, pour simplifier, nous n'exploiterons pas cette possibilité et nous prendrons $\alpha = 0$ et $\gamma = \frac{1}{2}$.

Remarque 3.3 $(A_- \cup A_+)^c = \{\exists t \in [a,b]\ t.q.\ \langle X_t\,,\, u\rangle = 0\}$.

Montrons le résultat suivant :

Théorème 3.4 *Il existe $k > 0$, $\varepsilon_0 > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,*

$$P\left((A_- \cup A_+)^c \,|\, C^\epsilon\right) \leq e^{-k/\sqrt{\epsilon}}.$$

Preuve Il existe $\theta > 0$, tel que:

$$\begin{aligned}(A_- \cup A_+)^c \cap C^\epsilon &\subset \left\{\sup_{[a,b]} |\langle H_+^{-1} h(X_t) - \hat{X}_t^+ , u\rangle| > \theta\right\} \\ &\subset \left\{\sup_{[a,b]} |h(X_t) - H_+ \hat{X}_t^+| > \theta |H_+|\right\}.\end{aligned}$$

D'après la proposition 3.1, il reste à démontrer que

$$\liminf_{\epsilon \to 0} P(C^\epsilon) > 0$$

Introduisons les événements suivants :

$$\begin{aligned}A_r &= \{|\langle X_t , u\rangle| > r , \forall t \in [a,b]\}, \\ D_\rho &= \left\{\sup_{[a,b]} |h(X_t) - H_+ \hat{X}_t^+| \leq \rho\right\}.\end{aligned}$$

Pour $\rho > 0$ et $r > 0$ convenablement choisis, on a l'inclusion $A_r \cap D_\rho \subset C^\epsilon$.

Montrons que $P(A_r) > 0$. Pour cela, plaçons nous sous les conditions d'application du théorème du support d'une diffusion de Stroock-Varadhan (cf. [7]) i.e. $\sigma \in C_b^2(\mathbb{R}^n)$ et $b \in C_b^1(\mathbb{R}^n)$.

Soit $0 < a_0 < a < b$. La mesure de Lebesgue de Δ_u étant nulle et la loi de X_t admettant une densité, on a donc:

$$P(X_{a_0} \in \Delta_u) = 0.$$

Il existe $\theta > 0$ tel que $P(\{|\langle X_{a_0}, u\rangle| > \theta\}) > 0$.

Traitons par exemple le cas où $P(\{\langle X_{a_0}, u\rangle > \theta\}) > 0$. Nous introduisons le temps d'arrêt τ,

$$\tau = \inf\{t \geq a_0 , \langle X_t , u\rangle = 0\}$$

Pour tout $r > 0$, on a:

$$\begin{aligned}&P\left(\{\langle X_t , u\rangle > r , \forall t \in [a,b]\} \mid \{\langle X_{a_0}, u\rangle > \theta\}\right) \\ \geq\ &P\left(\{\langle X_t , u\rangle > r , \forall t \in [a,b]\} \cap \{\tau \geq a\} \mid \{\langle X_{a_0}, u\rangle > \theta\}\right) > 0\end{aligned}$$

d'après le théorème du support. A partir de l'égalité $P(A_r \cap D_\rho) = P(A_r) - P(A_r \cap D_\rho^c)$ et de la proposition 3.1, on déduit que $\liminf_{\epsilon \to 0} P(C^\epsilon) \geq P(A_r) > 0$.

□

Par la suite, l'intervalle $[a,b]$ représentera un intervalle de temps fixe durant lequel la sortie du filtre de Kalman-Bucy (FK_+) verifie $C^\epsilon(a,b)$, et en conséquence la probabilité conditionnelle $P(A_-(a,b) \cup A_+(a,b) \mid \mathcal{Y}_b)$ est proche de 1. La prochaine étape consiste à choisir entre les deux alternatives "A_-" et "A_+".

4 Sous $(HD1)$, le test de la variation quadratique

Nous présentons un premier test, dit de la variation quadratique. Sous l'hypothèse de dissymètrie $(HD1)$: $H_-\Sigma_-\Sigma_-^*H_-^* \neq H_+\Sigma_+\Sigma_+^*H_+^*$, ce test permet de décider du signe de $\langle X_t\,,\,u\rangle$ avec une faible probabilité d'erreur. L'idée est la suivante: si $\varepsilon = 0$, $h(X_t)$ est observé exactement et nous pouvons décider du signe de $\langle X_t\,,\,u\rangle$ en calculant la variation quadratique de $dY_t/dt = h(X_t)$. Pour $\varepsilon > 0$, $h(X_t)$ ne correspond plus à l'observation, mais, pour ε petit, on peut calculer de façon approchée la variation quadratique d'une approximation de $h(X_t)$. Nous montrons que nous pouvons ainsi décider entre "A_-" et "A_+" avec une faible probabilité d'erreur. Ce test a été proposé par W.A. Fleming et E. Pardoux [6] pour un modèle plus simple que le nôtre. Leurs résultats s'adaptant sans difficulté à notre modèle, nous ne donnerons que les grandes lignes des démonstrations.

Soit $0 < a < b$, $m = [\frac{b-a}{\varepsilon}]$, on introduit les notations suivantes : pour $k = 0, \cdots, m-1$, $t_k = a + k\varepsilon$,

$$\begin{aligned} \bar{y}_k &= \varepsilon^{-1}\,(Y_{t_{k+1}} - Y_{t_k}), \\ s_k &= \varepsilon^{-1}\int_{t_k}^{t_{k+1}} h(X_s)\,ds, \\ \bar{w}_k &= W_{t_{k+1}} - W_{t_k}. \end{aligned}$$

On note que $\bar{y}_k = s_k + \bar{w}_k$.

Remarque 4.1 *La matrice de covariance de $\bar{w}_k$ est $\varepsilon\, I_n$.*

Pour $0 < \varepsilon \ll 1$, on approche $h(X_t)$, $t \in [t_k, t_{k+1}]$, par $\bar{y}_k$.
On pose

$$Z_i = \frac{1}{b-a}\sum_{k=0,k\,impair}^{m-2} (\bar{y}_{k+1} - \bar{y}_k)(\bar{y}_{k+1} - \bar{y}_k)^*,$$

$$Z_p = \frac{1}{b-a}\sum_{k=0,k\,pair}^{m-2} (\bar{y}_{k+1} - \bar{y}_k)(\bar{y}_{k+1} - \bar{y}_k)^*.$$

Sur A_+ :

$$Z_i = \frac{1}{b-a} \sum_{k=0, k\, impair}^{m-2} (\alpha_k \alpha_k^* + \beta_k \beta_k^* + \alpha_k \beta_k^* + \beta_k \alpha_k^*)$$

avec

$$\alpha_k = \frac{1}{\varepsilon} H_+ \int_{t_k}^{t_{k+1}} [F_+(V_{s+\epsilon} - V_s) + G_+(W_{s+\epsilon} - W_s)] ds + \bar{w}_{k+1} - \bar{w}_k,$$

$$\beta_k = \frac{1}{\varepsilon} H_+ B_+ \int_{t_k}^{t_{k+1}} \int_s^{s+\epsilon} X_u \, du \, ds.$$

On remarque que le vecteur aléatoire α_k suit une loi normale $N(O, 2\varepsilon(I_n + \frac{1}{3}\Gamma_+\Gamma_+^*))$ avec $\Gamma_+\Gamma_+^* = H_+\Sigma_+\Sigma_+^* H_+^*$ et que les sous suites $\{\alpha_k, k\, pair\}$, $\{\alpha_k, k\, impair\}$ sont des sous suites de variables aléatoires à valeurs dans $\mathbb{R}^n$, indépendantes et identiquement distribuées, contrairement à toute la suite $\{\alpha_k\}$.

Quand $\varepsilon \to 0$, on montre les convergences suivantes en probabilité:

Sur A_- :

$$Z_i \to I_n + \frac{1}{3}\Gamma_-\Gamma_-^*,$$

$$Z_p \to I_n + \frac{1}{3}\Gamma_-\Gamma_-^*.$$

Sur A_+ :

$$Z_i \to I_n + \frac{1}{3}\Gamma_+\Gamma_+^*,$$

$$Z_p \to I_n + \frac{1}{3}\Gamma_+\Gamma_+^*.$$

Ainsi sur $C^e(a,b)$, sous l'hypothèse $(HD1)$, il sera possible de choisir entre $A_-(a,b)$ et $A_+(a,b)$ au vu de Z_i ou Z_p .

Lemme 4.2 *$\forall \theta > 0$, il existe $k > 0$, $\varepsilon_0 > 0$ tels que: $\forall \epsilon \in (0, \varepsilon_0]$,*

$$P\left(\left\{|Z_i - (I_n + \frac{1}{3}\Gamma_+\Gamma_+^*)| \geq \theta\right\} \cap A_+\right) \leq e^{-k/\sqrt{\epsilon}},$$

$$P\left(\left\{|Z_i - (I_n + \frac{1}{3}\Gamma_-\Gamma_-^*)| \geq \theta\right\} \cap A_-\right) \leq e^{-k/\sqrt{\epsilon}}.$$

De même avec Z_p.

Preuve (cf.W.H. Fleming–E. Pardoux [6][lemme 4.1]).

Présentons le raisonnement pour la première inégalité : Sur A_+, trois événements sont considérés :

$$G = \left\{ \left| \frac{2\varepsilon}{b-a} \sum_{k=0,k\,impair}^{m-2} \frac{1}{\varepsilon}[\alpha_k \alpha_k^*] - \frac{1}{\varepsilon}\mathbf{E}(\alpha_1 \alpha_1^*) \right| > \frac{2\theta}{3} \right\},$$

$$H = \left\{ \left| \sum_{k=0,k\,impair}^{m-2} \beta_k \beta_k^* \right| > \frac{b-a}{3}\theta \right\},$$

$$J = \left\{ \left| \sum_{k=0,k\,impair}^{m-2} (\alpha_k \beta_k^* + \beta_k \alpha_k^*) \right| > \frac{b-a}{3}\theta \right\}.$$

On montre que les probabilités de ces trois événements sont exponentiellement petite quand $\varepsilon \to 0$, en utilisant respectivement, un résultat de grandes déviations (cf.Ellis [3][th II.3.3-II.4.1]), le lemme 2.7 et le caractère gaussien des variables indépendantes et identiquement distribuées de la suite $\{\alpha_k, k\,impair\}$. On conclut en remarquant que

$$\mathbf{E}(\alpha_1 \alpha_1^*) = 2\varepsilon \left(I_n + \frac{1}{3}\Gamma_+ \Gamma_+^* \right).$$

□

Construisons le test de la variation quadratique:

On définit deux événements tests sur $C^\varepsilon(a,b)$:

$$C_-^\varepsilon(a,b) = C^\varepsilon(a,b) \cap \left\{ \left| Z_i - (I_n + \frac{1}{3}\Gamma_- \Gamma_-^*) \right| \leq \left| Z_i - (I_n + \frac{1}{3}\Gamma_+ \Gamma_+^*) \right| \right\}$$

$$C_+^\varepsilon(a,b) = C^\varepsilon(a,b) \cap \left\{ \left| Z_i - (I_n + \frac{1}{3}\Gamma_- \Gamma_-^*) \right| > \left| Z_i - (I_n + \frac{1}{3}\Gamma_+ \Gamma_+^*) \right| \right\}$$

Remarque 4.3 $C_+^\varepsilon \cup C_-^\varepsilon = C^\varepsilon$ et $C_+^\varepsilon \cap C_-^\varepsilon = \emptyset$.

Lemme 4.4 *Il existe* $k > 0$, $\varepsilon_0 > 0$ *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P(A_-^c \cap C_-^\varepsilon) \leq e^{-k/\sqrt{\varepsilon}},$$
$$P(A_+^c \cap C_+^\varepsilon) \leq e^{-k/\sqrt{\varepsilon}}.$$

Preuve En utilisant l'inégalité suivante:

$$P(A_+^c \cap C_+^\epsilon) \leq P((A_- \cup A_+)^c \cap C^\epsilon) + P\left(A_- \cap \left\{\left|Z_i - (I_n + \frac{1}{3}\Gamma_-\Gamma_-^*)\right| \geq \frac{1}{6}\left|\Gamma_+\Gamma_+^* - \Gamma_-\Gamma_-^*\right|\right\}\right),$$

le lemme découle du théorème 3.4 et du lemme 4.2.

□

Procédure du test de la variation quadratique :

Soit $\tau_1 > 0$, $a < \tau_1 < b$. Sur $C^\epsilon(a,b)$: on décide entre "$A_+(a,b)$" et "$A_-(a,b)$" au vu de l'estimation de la variation quadratique calculée à partir des observations recuellies durant l'intervalle de temps $[a,\tau_1]$, avec $a < \tau_1 \leq b$, τ_1 représentant le temps à partir duquel on peut prendre un décision. On introduit deux nouveaux événements tests:

$$Q_-^\epsilon(a,b) = C^\epsilon(a,b) \cap C_-^\epsilon(a,\tau_1), \quad (4.1)$$
$$Q_+^\epsilon(a,b) = C^\epsilon(a,b) \cap C_+^\epsilon(a,\tau_1). \quad (4.2)$$

On montre à l'aide du lemme 4.4 et du théorème 3.4 le résultat suivant :

Théorème 4.5 *Il existe $k > 0$, $\epsilon_0 > 0$ tels que $\forall \epsilon \in (0,\epsilon_0]$:*

$$P(A_-(a,b)^c \,|\, Q_-^\epsilon(a,b)) \leq e^{-k/\sqrt{\epsilon}}$$
$$P(A_+(a,b)^c \,|\, Q_+^\epsilon(a,b)) \leq e^{-k/\sqrt{\epsilon}}$$

$$P((A_-(a,b) \cup A_+(a,b))^c \cap (Q_-^\epsilon(a,b) \cup Q_+^\epsilon(a,b)) \leq e^{-k/\sqrt{\epsilon}}.$$

□

Remarque 4.6 Ce test a été mis en œuvre par W.H. Fleming, D. Ji, P. Salame et Q. Zhang [5] de deux façons:

- test à taille d'échantillon fixée: la durée $\tau_1 - a$ est fixée au préalable,
- test séquentiel: la durée $\tau_1 - a$ est aléatoire.

Leur étude montre que pour une même probalilité d'erreur, le second type de test permet de décider plus rapidement que le premier.

5 Sous $(HD2)$, le test du rapport de vraisemblance sur les accroissements de l'observation

Nous nous intéressons au cas où $(HD1)$ n'est pas vérifiée et en conséquence le test de la variation quadratique ne peut être appliqué. Sous l'hypothèse

$$(HD2) : \begin{cases} \Gamma\Gamma^* = \Gamma_-\Gamma_-^* = \Gamma_+\Gamma_+^*, \\ \ker\{ H_-B_-H_-^{-1} - H_+B_+H_+^{-1} \} \subset \Delta_u \\ \text{la matrice } (\Gamma\Gamma^*)^{-1}(H_-B_-H_-^{-1} - H_+B_+H_+^{-1}) \text{ est symétrique,} \end{cases}$$

nous proposons un test du type rapport de vraisemblance, basé sur les accroissements du processus d'observation. Ce test est inspiré du cas limite $\varepsilon = 0$.

Remarque 5.1 Dans le paragraphe 6, nous présenterons un test ne nécessitant pas la symétrie de la matrice $(\Gamma\Gamma^*)^{-1}(H_-B_-H_-^{-1} - H_+B_+H_+^{-1})$. Dans le cas où $n = 1$, l'hypothèse $(HD2)$ devient:

$$(HD2) : \quad \begin{cases} \Gamma_+^2 = \Gamma_-^2 \\ B_+ \neq B_- \end{cases}$$

5.1 Cas $\varepsilon = 0$

Sur l'espace de probabilité $(\Omega^1, \mathcal{F}^1, \mathcal{F}_t^1, P^X)$ défini dans le paragraphe 2, on considère le système suivant:

$$\begin{cases} dX_t &= b(X_t)\,dt + \sigma(X_t)dU_t, \\ y_t &= h(X_t). \end{cases}$$

Par abus de notation, nous désignerons par $A_-(a,b)$ et $A_+(a,b)$ les traces de ces événements sur Ω^1. On pose $\mathcal{F}_t^{1,a} = \sigma\{X_s - X_a\,,\, a \leq s \leq t\}$

Notre problème est de déterminer le signe de $\langle X_t\,,\, u\rangle$ pour $t \in [a,b]$. D'après l'hypothèse $(H6)$, on a l'équivalence suivante,

$$\langle X_t\,,\, u\rangle \neq 0 \Longleftrightarrow \langle y_t\,,\, u\rangle \neq 0.$$

Si $\langle y_t\,,\, u\rangle \neq 0$ pour $t \in [a,b]$, nous pouvons proposer un test d'hypothèses sur les observations $\{y_t\,;\, a \leq t \leq b\}$.

Sur $A_+(a,b)$:

$$dy_t = (H_+ B_+ H_+^{-1})\, y_t\, dt + \Gamma dW_t^+, \tag{5.1}$$

avec $\{W_t^+ - W_a^+ \,;\, a \le t \le b\}$ un processus de Wiener standard sous P^+, la loi de probabilité définie sur $(\Omega^1, \mathcal{F}_b^{1,a})$ solution du problème de martingales associé à (5.1).

Sur $A_-(a,b)$:

$$dy_t = (H_- B_- H_-^{-1})\, y_t\, dt + \Gamma dW_t^-, \tag{5.2}$$

avec $\{W_t^- - W_a^- \,;\, a \le t \le b\}$ un processus de Wiener standard sous P^-, la loi de probabilité définie sur $(\Omega^1, \mathcal{F}_b^{1,a})$ solution du problème de martingales associé à (5.2).

Remarque 5.2

$$\begin{aligned} P^X(D) &= P^+(D) \text{ si } D \subset A_+(a,b), \\ P^X(D) &= P^-(D) \text{ si } D \subset A_-(a,b). \end{aligned}$$

Remarque 5.3 Notons M_t la partie martingale de y_t. Sur l'événement $A_- \cup A_+$, on a:

$$\begin{aligned} \langle M_i\,,\, M_j \rangle_t &= \Gamma_{ii}^2\, t \qquad \text{si } i = j, \\ \langle M_i\,,\, M_j \rangle_t &= 0 \qquad\quad \text{si } i \neq j. \end{aligned}$$

Le fait que la matrice Γ soit inversible nous permet de construire une statistique de test du rapport de vraisemblance. D'après le théorème de Girsanov, $L(a,b) = \left(\frac{dP^+}{dP^-}\right)$ est donnée par

$$\begin{aligned} L(a,b) = & \int_a^b y_s^* (H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1})^* (\Gamma\Gamma^*)^{-1}\, dy_s \\ & - \frac{1}{2} \int_a^b \left(|\Gamma^{-1}(H_+ B_+ H_+^{-1})\, y_s|^2 - |\Gamma^{-1}(H_- B_- H_-^{-1})\, y_s|^2 \right) ds. \end{aligned} \tag{5.3}$$

Remarque 5.4 En dimension 1, les expressions se simplifient (voir [10]). La statistique $L(a,b)$ s'écrit:

$$L(a,b) = \frac{1}{\Gamma^2}\left[(B_+ - B_-)\int_a^b y_s\, dy_s - \frac{1}{2}\int_a^b \left((B_+ y_s)^2 - (B_- y_s)^2\right) ds\right].$$

Pour $a \leq t \leq b$, on définit:

$$\begin{aligned} M_t^+ &= \int_a^t y_s^*(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1})^*(\Gamma^*)^{-1} dW_s^+, \\ M_t^- &= \int_a^t y_s^*(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1})^*(\Gamma^*)^{-1} dW_s^-. \end{aligned}$$

Réécrivons la statistique $L(a,b)$ respectivement sous la loi P^+ et la loi P^-:

$$\begin{aligned} L(a,b) &= \frac{1}{2}\langle M^+\rangle_b + M_b^+, \\ L(a,b) &= -\frac{1}{2}\langle M^-\rangle_b + M_b^-. \end{aligned}$$

On remarque que $\langle M^+\rangle_b = \langle M^-\rangle_b$ et on introduit la notation $R(a,b)$ pour désigner ces variations quadratiques. Sur $\{\langle y_t\,,\, u\rangle \neq 0\,;\, a \leq t \leq b\}$, on utilise la règle de décision suivante:

- Si $L(a,b) \geq 0$: on décide "A_+",
- Si $L(a,b) < 0$: on décide "A_-".

La qualité de la décision dépend de la quantité $R(a,b)$:

$$R(a,b) = \int_a^b |\Gamma^{-1}(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1})\, y_s|^2\, ds. \qquad (5.4)$$

Montrons que pour R assez grand, la probabilité d'erreur est faible.

Proposition 5.5 *Soit* $r > 0$,

$$\begin{aligned} P^X(\{L < 0\} \cap \{R \geq r\} \cap A_+) &\leq e^{-r/8}, \\ P^X(\{L \geq 0\} \cap \{R \geq r\} \cap A_-) &\leq e^{-r/8}. \end{aligned}$$

Preuve On démontre la premiére inégalité. D'après la remarque 5.2, on a

$$P^X(\{L<0\}\cap\{R\geq r\}\cap A_+) = P^+(\{L<0\}\cap\{R\geq r\}\cap A_+)$$

Du lemme 2.7 on déduit l'existence de $k>0$ tel que:

$$\mathbf{E}^+\left(\exp\left\{k\sup_{[a,b]}|X_t|^2\right\}\right)<\infty. \tag{5.5}$$

On a donc

$$\mathbf{E}^+\left(\exp\left\{\mu M_t^+ - \frac{\mu^2}{2}\langle M^+\rangle_t\right\}\right)=1\,,\ \mu\in\mathbb{R}. \tag{5.6}$$

Soit $-1<\mu<0$,

$$\begin{aligned}
& P^+(\{L<0\}\cap\{R>r\}\cap A_+)\\
\leq\ & P^+\left(\left\{-M_b^+>\frac{1}{2}R\right\}\cap\{R>r\}\right)\\
\leq\ & P^+\left(\exp\left\{\mu M_b^+ - \frac{\mu^2}{2}\langle M^+\rangle_b\right\} > \exp\left\{(-\frac{1}{2}\mu-\frac{\mu^2}{2})r\right\}\right)\\
\leq\ & \exp\left\{\frac{\mu}{2}(1+\mu)r\right\},
\end{aligned}$$

d'après (5.6) et l'inégalité de Markov. Nous obtenons le résultat en choisissant $\mu=-\frac{1}{2}$.

□

Montrons maintenant que pour R grand, la probabilité pour que la satistique L prenne des valeurs proches de zéro est faible.

Proposition 5.6 *Soit* $r>0$, $\forall\theta>0$, $\exists k_1>0$, $k_2>0$,

$$\begin{aligned}
P^X(\{L<\theta\}\cap\{R\geq r\}\cap A_+) &\leq e^{k_1\theta}e^{-k_2 r},\\
P^X(\{L>-\theta\}\cap\{R\geq r\}\cap A_-) &\leq e^{k_1\theta}e^{-k_2 r}.
\end{aligned}$$

Preuve On considère uniquement la première probabilité.

$$\begin{aligned}
& P^X(\{L<\theta\}\cap\{R\geq r\}\cap A_+)\\
\leq\ & P^+\left(\left\{M_b^+ + \frac{1}{2}R<\theta\right\}\cap\{R\geq r\}\right).
\end{aligned}$$

On pose

$$Z_t = \exp\left\{ \mu M_t^+ - \frac{\mu^2}{2}\langle M^+\rangle_t \right\}$$

avec $-1 < \mu < 0$. D'après (5.6), $\{Z_t, a \le t \le b\}$ est une P^+-martingale et $\mathbf{E}^+(Z_t) = 1$. On a

$$\begin{aligned} & P^+\left(\left\{M_b^+ + \frac{1}{2}R < \theta\right\} \cap \{R \ge r\}\right) \\ \le\ & P^+\left(\left\{\mu M_b^+ - \frac{\mu^2}{2}R > \mu\theta - \frac{\mu}{2}(1+\mu)r\right\}\right) \\ \le\ & P^+\left(\left\{Z_b > \exp\{\mu\theta - \frac{\mu}{2}(1+\mu)r\right\}\right) \\ \le\ & e^{-\mu\theta}e^{\frac{\mu}{2}(1+\mu)r} \le e^{k_1\theta}e^{-k_2 r}, \end{aligned}$$

avec $k_1 > 0$, $k_2 > 0$.

□

En préparation du cas $\varepsilon > 0$, sur l'événement $\{y_t \neq 0, a \le t \le b\}$, on réécrit L comme une fonction continue de la trajectoire $y_t = h(X_t)$ en utilisant la formule d'Itô et la remarque 5.3. Pour cela, nous définissons la fonction suivante:

$$\begin{aligned} \kappa:\ & \mathbb{R}^n \ \to \mathbb{R} \\ \kappa:\ & x \ \to \tfrac{1}{2}x^*(\Gamma\Gamma^*)^{-1}(H_+B_+H_+^{-1} - H_-B_-H_-^{-1})x \end{aligned}$$

$$\begin{aligned} L = \quad & \kappa(h(X_b)) - \kappa(h(X_a)) \\ - \ & \frac{1}{2}(b-a)\ \text{Trace}\ (H_+B_+H_+^{-1} - H_-B_-H_-^{-1}) \\ - \ & \frac{1}{2}\int_a^b \left(|\Gamma^{-1}(H_+B_+H_+^{-1})\,h(X_s)|^2 - |\Gamma^{-1}(H_-B_-H_-^{-1})\,h(X_s)|^2\right) ds. \end{aligned} \tag{5.7}$$

5.2 Cas $\varepsilon > 0$

On utilise les notations introduites dans le paragraphe 3. Pour ε petit, comme pour le test de la variation quadratique, on approche $h(X_t)$ par les

accroissements $\bar{y}_k$ de l'observation. On se place à nouveau dans l'espace de probabilité $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.

On note $\bar{L}$ (resp. $\bar{R}$) l'extension de L (resp. R) à Ω. $\forall \omega = (\omega_1, \omega_2) \in \Omega$, $\bar{L}(\omega_1, \omega_2) = L(\omega_1)$ et $\bar{R}(\omega_1, \omega_2) = R(\omega_1)$.

On définit:

$$\begin{aligned} L^\epsilon &= \kappa(\bar{y}_{m-1}) - \iota(\bar{y}_0) \\ &- \frac{1}{2}(b - a) \text{ Trace } (H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1}) \\ &- \frac{\epsilon}{2} \sum_{k=0}^{m-1} \left(|\Gamma^{-1}(H_+ B_+ H_+^{-1})\, \bar{y}_k|^2 - |\Gamma^{-1}(H_- B_- H_-^{-1})\, \bar{y}_k|^2 \right) \end{aligned}$$

et

$$R^\epsilon = \epsilon \sum_{k=0}^{m-1} |\Gamma^{-1}(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1})\, \bar{y}_k|^2.$$

On montre tout d'abord que sur C^ϵ, l'erreur d'approximation est faible avec une probabilité proche de 1 pour ϵ petit.

Lemme 5.7 $\forall \theta > 0$, $\forall \Lambda$ *matrice* (n,n) *symétrique, il existe* $\epsilon_0 > 0$, $k > 0$ *tels que* $\forall \epsilon \in (0, \epsilon_0]$,

$$P\left(C^\epsilon \cap \left\{ \max_{0 \le k \le m-1} \sup_{[t_k, t_{k+1}]} |h(X_s)^* \Lambda\, h(X_s) - \bar{y}_k^* \Lambda\, \bar{y}_k| > \theta \right\} \right) \le e^{-k/\sqrt{\epsilon}}.$$

Preuve D'après le théorème 3.4, il suffit de s'intéresser à la probabilité de l'événement:

$$A_+ \cap \left\{ \max_{0 \le k \le m-1} \sup_{[t_k, t_{k+1}]} |h(X_s)^* \Lambda\, h(X_s) - \bar{y}_k^* \Lambda\, \bar{y}_k| > \theta \right\}.$$

Sur l'événement A_-, le raisonnement est identique.

Sur A_+:

$$\begin{aligned}
&\sup_{[t_k,t_{k+1}]} |h(X_s)^* \Lambda\, h(X_s) - \bar{y}_k^* \Lambda\, \bar{y}_k| \\
= &\sup_{[t_k,t_{k+1}]} |X_s^* H_+^* \Lambda H_+ X_s - \bar{y}_k^* \Lambda\, \bar{y}_k| \\
\leq &\sup_{t_k \leq s < u \leq t_{k+1}} |X_s^* H_+^* \Lambda H_+ X_s - X_u^* H_+^* \Lambda H_+ X_u| + |\Lambda|\,|\bar{w}_k|^2 + 2\,|\Lambda|\,|\bar{w}_k|\,|s_k|.
\end{aligned}$$

D'autre part, l'expression $X_t^* H_+^* \Lambda H_+ X_t$ s'écrit:

$$\begin{aligned}
X_t^* H_+^* \Lambda H_+ X_t \;=\; & X_a^* H_+^* \Lambda\, H_+ X_a + t \text{ Trace } (H_+^* \Lambda H_+ \Sigma_+ \Sigma_+^*) \\
& + 2 \int_a^t (H_+^* \Lambda H_+ X_s)^* \, dX_s.
\end{aligned}$$

D'où

$$\begin{aligned}
&\sup_{t_k \leq s < u \leq t_{k+1}} |X_s^* H_+^* \Lambda H_+ X_s - X_u^* H_+^* \Lambda H_+ X_u| \\
\leq \quad & \epsilon \text{ Trace } (H_+^* \Lambda H_+ \Sigma_+ \Sigma_+^*) \\
+ \quad & \sup_{t_k \leq s < u \leq t_{k+1}} \left| 2 \int_s^u X_r^* H_+^* \Lambda H_+ B_+ X_r \, dr \right| \\
+ \quad & \sup_{t_k \leq s < u \leq t_{k+1}} \left| 2 \int_s^u X_r^* H_+^* \Lambda H_+ \Sigma_+ \, dU_r \right|.
\end{aligned}$$

Pour ϵ suffisament petit, nous obtenons:

$$\begin{aligned}
& P\left(A_+ \cap \left\{ \max_{0 \leq k \leq m-1} \sup_{[t_k,t_{k+1}]} |h(X_s)^* \Lambda h(X_s) - \bar{y}_k^* \Lambda \bar{y}_k| > \theta \right\} \right) \\
\leq \quad & P\left(\left\{ \sup_{[a,b]} |X_s|^2 > \theta_1/\epsilon \right\} \right) \\
+ \quad & P\left(\left\{ \max_{0 \leq k \leq m-1} \sup_{t_k \leq s < u \leq t_{k+1}} \left| \int_s^u X_r^* H_+^* \Lambda H_+ \Sigma_+ \, dU_r \right| > \theta_2 \right\} \right) \\
+ \quad & P\left(\left\{ \max_{0 \leq k \leq m-1} |\bar{w}_k^2| > \theta_3 \right\} \right) + P\left(A_+ \cap \left\{ \max_{0 \leq k \leq m-1} |\bar{w}_k||s_k| > \theta_4 \right\} \right),
\end{aligned}$$

avec θ_1, θ_2, θ_3 et $\theta_4 > 0$ choisis de façon adéquate. On s'intéresse à chacun des quatre termes de droite de l'inégalité.

1ier terme: D'après le lemme 2.7, il existe $\eta > 0$ tel que

$$\mathbf{E}\left(\exp\eta\left\{\sup_{[a,b]}|X_s|^2\right\}\right) < \infty.$$

Par l'inégalité de Markov, on obtient l'existence de $k_1 > 0$, $\varepsilon_1 > 0$, tels que $\forall \varepsilon \in (0, \varepsilon_1]$,

$$P\left(\left\{\sup_{[a,b]}|X_s|^2 > \theta_1/\varepsilon\right\}\right) \leq e^{-k_1/\varepsilon}. \tag{5.8}$$

2ième terme: On pose $M_t = \int_0^t X_s^* H_+^* \Lambda H_+ \Sigma_+ dU_s$.

$$P\left(\left\{\max_{0\leq k\leq m-1}\sup_{t_k\leq s<u\leq t_{k+1}}|M_u - M_s| > \theta_2\right\}\right)$$
$$\leq \sum_{k=0}^{m-1} P\left(\left\{\sup_{[t_k,t_{k+1}]}|M_s - M_{t_k}| > \theta_2/2\right\}\right).$$

Soit $Z_t = \exp\{\mu M_t - \frac{\mu}{2}\langle M\rangle_t\}$, $\mu \in \mathbb{R}^*$. Prennons $\mu = \frac{1}{\sqrt{\varepsilon}}$. D'après le lemme 2.7, le processus $\{Z_t\,,\, t \geq 0\}$ est une $\mathcal{F}_t$-martingale telle que

$$\mathbf{E}(Z_t) = 1 \quad \forall t \geq 0.$$

De plus, $\forall p \geq 1$, il existe $\varepsilon_p > 0$ tel que $\forall \varepsilon \in (0, \varepsilon_p]$, $\{Z_s/Z_{t_k}\,,\, s \in [t_k, t_{k+1}]\}$ est une $\mathcal{F}_s$-martingale de puissance p – ième intégrable.

Pour $k = 0, \cdots, m-1$, on définit $Z_s^k = Z_s/Z_{t_k}$ et l'événement D_k,

$$D_k = \left\{\langle M\rangle_{t_{k+1}} - \langle M\rangle_{t_k} > \theta_2\sqrt{\varepsilon}/2\right\}$$

On remarque que

$$P(D_k) < P\left(\left\{\sup_{[t_k,t_{k+1}]}|X_s|^2 > \frac{\theta_2}{2|H_+^*\Lambda H_+\Sigma_+|^2\sqrt{\varepsilon}}\right\}\right)$$

et du lemme 2.7, on obtient

$$P(D_k) \leq e^{-k/\sqrt{\varepsilon}},$$

pour un certain $k > 0$ et ε assez petit. Il suffit donc de s'intéresser à la probabilité de l'événement $\{\sup_{[t_k,t_{k+1}]} |M_s - M_{t_k}| > \theta_2/2\} \cap D_k^c$. Or,

$$P\left(\left\{\sup_{[t_k,t_{k+1}]}(M_s - M_{t_k}) > \theta_2/2\right\} \cap D_k^c\right)$$
$$\leq P\left(\left\{\sup_{[t_k,t_{k+1}]} Z_s^k > e^{\theta_2/4\sqrt{\varepsilon}}\right\}\right) \leq k e^{-\theta_2/4\sqrt{\varepsilon}},$$

d'après l'inégalité de Doob.

En choisissant $\mu = -\frac{1}{\sqrt{\varepsilon}}$, on montre de même que

$$P\left(\left\{\sup_{[t_k,t_{k+1}]} -(M_s - M_{t_k}) > \theta_2/2\right\} \cap D_k^c\right) \leq k e^{-\theta_2/4\sqrt{\varepsilon}}.$$

Il existe donc $k_2 > 0$, $\varepsilon_2 > 0$, tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\max_{0\leq k\leq m-1} \sup_{t_k \leq s < u \leq t_{k+1}} |M_u - M_s| > \theta_2\right\}\right) \leq e^{-k_2/\sqrt{\varepsilon}}. \tag{5.9}$$

$3^{\text{ième}}$ **terme:** On s'intéresse à $\{\max_{0\leq k\leq m-1} |\bar{w}_k|^2 > \theta_3\}$. Nous savons que les variables aléatoires $\bar{w}_0, \cdots, \bar{w}_{m-1}$, à valeurs dans $\mathbb{R}^n$ sont indépendantes et de loi commune $N(0, \varepsilon I_n)$. Le lemme 2.4 entraine le résultat suivant:

$$P\left(\left\{\max_{0\leq k\leq m-1} |\bar{w}_k|^2 > \theta_3\right\}\right) \leq n m e^{-\theta_3/2\varepsilon n^2}. \tag{5.10}$$

$4^{\text{ième}}$ **terme:**

$$P\left(\left\{\max_{0\leq k\leq m-1} |\bar{w}_k|\,|s_k| > \theta_4\right\} \cap A_+\right)$$
$$\leq P\left(\left\{\sup_{[a,b]} |X_s| > \varepsilon^{-1/4}\frac{\sqrt{\theta_4}}{|H_+|}\right\}\right) + P\left(\left\{\max_{0\leq k\leq m-1} |\bar{w}_k| > \varepsilon^{1/4}\sqrt{\theta_4}\right\}\right)$$
$$\leq e^{-k_4/\sqrt{\varepsilon}}, \tag{5.11}$$

avec $k_4 > 0$ et ε assez petit, d'après les lemmes 2.7 et 2.4.

Les inégalités (5.8), (5.9), (5.10) et (5.11) permettent de conclure.

□

De ce résultat, on déduit que les différences $|R^\epsilon - \bar{R}|$ et $|L^\epsilon - \bar{L}|$ sont très petites avec une probabilité proche de 1.

Lemme 5.8 $\forall \theta > 0$, *il existe* ϵ_0, $k > 0$ *tels que* $\forall \epsilon \in (0, \epsilon_0]$,

$$P(C^\epsilon \cap \{|R^\epsilon - \bar{R}| > \theta\}) \leq e^{-k/\sqrt{\epsilon}}.$$

□

Lemme 5.9 $\forall \theta > 0$, *il existe* ϵ_0, $k > 0$ *tels que* $\forall \epsilon \in (0, \epsilon_0]$,

$$P(C^\epsilon \cap \{|L^\epsilon - \bar{L}| > \theta\}) \leq e^{-k/\sqrt{\epsilon}}.$$

□

On choisit arbitrairement $r > 0$ et on définit les événements suivants:

$$\begin{aligned} C^{\epsilon,r}(a,b) &= C^\epsilon(a,b) \cap \{R^\epsilon \geq r\}, \\ C^{\epsilon,r}_+(a,b) &= C^{\epsilon,r}(a,b) \cap \{L^\epsilon \geq 0\}, \\ C^{\epsilon,r}_-(a,b) &= C^{\epsilon,r}(a,b) \cap \{L^\epsilon < 0\}. \end{aligned}$$

Montrons que sur $C^{\epsilon,r}$, la statistique L^ϵ nous permet de choisir entre "A_+" et "A_-", avec une probabilité d'erreur dépendante de ε, r et $b - a$.

Proposition 5.10 *Soit* $r > 0$, *Il existe* $\varepsilon_0 > 0$, $k_1 > 0$, $k_2 > 0$, $k_3 > 0$ *tels que* $\forall \epsilon \in (0, \varepsilon_0]$,

$$\begin{aligned} P(A^c_+ \cap C^{\epsilon,r}_+) &\leq e^{-k_1/\sqrt{\epsilon}} + k_2\, e^{-k_3 r}, \\ P(A^c_- \cap C^{\epsilon,r}_-) &\leq e^{-k_1/\sqrt{\epsilon}} + k_2\, e^{-k_3 r}, \end{aligned}$$

Preuve On démontre la première inégalité:

$$A^c_+ \cap C^{\epsilon,r}_+ \subset (A_- \cap C^{\epsilon,r}_+) \cup (A_- \cup A_+)^c \cap C^\epsilon.$$

D'après le théorème 3.4, il suffit de s'intéresser à l'événement $A_- \cap C^{\epsilon,r}_+$. Il existe $\theta > 0$ tel que

$$\begin{aligned} A_- \cap C^{\epsilon,r}_+ \subset \quad & A_- \cap \{\bar{R} \geq r - \theta\} \cap \{\bar{L} \geq 0\} \\ \cup \; & A_- \cap \{\bar{R} \geq r - \theta\} \cap \{-\theta < \bar{L} < 0\} \\ \cup \; & (C^\epsilon \cap \{|\bar{R} - R^\epsilon| > \theta\}) \cup (C^\epsilon \cap \{|\bar{L} - L^\epsilon| > \theta\}) \end{aligned}$$

On conclut en utilisant les propositions 5.5-5.6 et les lemmes 5.8-5.9.

□

Intéressons nous au terme de précision R^ϵ. On définit une fonction $\delta_{[a,b]}$: $\mathbb{R}^+ \to [0,1[$ de la façon suivante:

$$\delta_{[a,b]}(x) = P\left(\left\{\inf_{[a,b]} |\Gamma^{-1}(H_+B_+H_+^{-1} - H_-B_-H_-^{-1})\, H_+\hat{X}_s^+|^2 \leq x\right\} \cap C^\epsilon(a,b)\right).$$

La fonction $\delta_{[a,b]}$ est croissante, bornée par $P(C^\epsilon(a,b))$ et il existe $c_{inf} > 0$ tel que $\forall x \in [0, c_{inf}]$, on a

$$\delta_{[a,b]}(x) = 0.$$

En effet, sur l'événement C^ϵ,

$$|\langle \hat{X}_s^+ , u\rangle| > c, \quad \forall s \in [a,b].$$

Donc, d'après l'hypothèse $(HD2)$ et le fait que $h(\Delta_u) = \Delta_u$, il existe $c_{inf} > 0$ tel que

$$\inf_{[a,b]} |\Gamma^{-1}(H_+B_+H_+^{-1} - H_-B_-H_-^{-1})\, H_+\hat{X}_s^+|^2 > c_{inf}. \tag{5.12}$$

On note que c_{inf} ne dépend pas de l'intervalle de temps $[a,b]$.

Lemme 5.11 *Soit $r > 0$, $\lambda > 1$, il existe $\varepsilon_0 > 0$, $k > 0$, tels que $\forall \epsilon \in (0, \varepsilon_0]$,*

$$P(\{R^\epsilon < r\} \cap C^\epsilon) \leq \delta_{[a,b]}\left(\frac{\lambda r}{b-a}\right) + e^{-k/\sqrt{\epsilon}}.$$

Preuve $\forall \theta > 0$, on a:

$$\{R^\epsilon < r\} \cap C^\epsilon \subset [\,\{|R^\epsilon - \bar{R}| > \theta\} \cap C^\epsilon\,] \cup [\,\{\bar{R} \leq r + \theta\} \cap C^\epsilon\,].$$

Rappelons l'expression de $\bar{R}$,

$$\bar{R} = \int_a^b |\Gamma^{-1}(\, H_+B_+H_+^{-1} - H_-B_-H_-^{-1}\,)\, h(X_s)|^2 ds.$$

D'après le lemme 5.8,

$$P(\{|R^\epsilon - \bar{R}| > \theta\} \cap C^\epsilon) \leq e^{-k/\sqrt{\epsilon}},$$

D'autre part, il existe $\theta = \theta(\lambda) > 0$ tel que

$$\begin{aligned} & \{\bar{R} \le r + \theta\} \cap C^\epsilon \\ \subset \quad & \left\{ \int_a^b |\Gamma^{-1}(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1})(h(X_s) - H_+ \hat{X}_s^+)|^2 \, ds > \theta \right\} \\ \cup \quad & \left\{ \int_a^b |\Gamma^{-1}(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1}) H_+ \hat{X}_s^+|^2 \, ds \le \lambda r \right\} \cap C^\epsilon. \end{aligned}$$

La proposition 3.1 implique que pour tout $\theta > 0$, la probabilité du premier événement est exponentiellement faible quand ϵ est petit. Ainsi, il suffit de s'intéresser à la probabilité

$$P\left(\left\{ \int_a^b |\Gamma^{-1}(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1}) H_+ \hat{X}_s^+|^2 \, ds \le \lambda r \right\} \cap C^\epsilon \right).$$

On conclut en remarquant que

$$\begin{aligned} & P\left(\left\{ \int_a^b |\Gamma^{-1}(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1}) H_+ \hat{X}_s^+|^2 \, ds \le \lambda r \right\} \cap C^\epsilon \right) \\ \le \quad & P\left(\left\{ \inf_{[a,b]} |\Gamma^{-1}(H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1}) H_+ \hat{X}_s^+|^2 \le \frac{\lambda r}{b-a} \right\} \cap C^\epsilon \right) \\ \le \quad & \delta_{[a,b]}\left(\frac{\lambda r}{b-a} \right). \end{aligned}$$

□

Remarque 5.12 Sur C^ϵ, le terme de précision R^ϵ est minoré avec une probabilité proche de 1, par $r_{inf} < (b-a)\, c_{inf}$, c_{inf} défini en (5.12).

Dans le résultat suivant, on met en évidence que le choix de r est un compromis entre réduire la probabilité d'erreur du test et maintenir une probabilité de décision assez grande.

Théorème 5.13 *Soit $\lambda > 1$, $r_0 > 0$, il existe $\varepsilon_0 > 0$, $k_1 > 0$, $k_2 > 0$, $k_3 > 0$ et une fonction croissante $\rho^\epsilon_{[a,b]} : \mathbb{R}^+ \to \mathbb{R}$, tels que $\forall r \in (0, r_0]$, $\forall \epsilon \in (0, \varepsilon_0]$,*

$$\begin{aligned} P(A_+^c \,|\, C_+^{\epsilon,r}) \quad & \le \quad e^{-k_1/\sqrt{\epsilon}} + \rho^\epsilon_{[a,b]}\left(\frac{\lambda r}{b-a} \right) k_2 \, e^{-k_3 r} \\ P(A_-^c \,|\, C_-^{\epsilon,r}) \quad & \le \quad e^{-k_1/\sqrt{\epsilon}} + \rho^\epsilon_{[a,b]}\left(\frac{\lambda r}{b-a} \right) k_2 \, e^{-k_3 r} \end{aligned}$$

Preuve D'après la proposition 5.10, il nous reste à minorer les probabilités $P(C_+^{\epsilon,r}(a,b))$ et $P(C_-^{\epsilon,r}(a,b))$. Considérons la première:

$$P(C_+^{\epsilon,r}(a,b)) = P(\{L^\epsilon \geq 0\} \cap C^\epsilon) - P(\{L^\epsilon \geq 0\} \cap C^\epsilon \cap \{R^\epsilon < r\}).$$

Le lemme 5.11 implique l'existence de $k > 0$ et de $\varepsilon_0 > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P(\{L^\epsilon \geq 0\} \cap C^\epsilon \cap \{R^\epsilon < r\}) \leq \delta_{[a,b]}\left(\frac{\lambda r}{b-a}\right) + e^{-k/\sqrt{\epsilon}}. \tag{5.13}$$

D'autre part, par un raisonnement similaire à celui de la démonstration du théorème 3.4, on montre qu'il existe β, $0 < \beta < 1$, tel que

$$P(\{L^\epsilon \geq 0\} \cap C^\epsilon) > \beta. \tag{5.14}$$

Les inégalités (5.13) et (5.14) impliquent:

$$P(C_+^{\epsilon,r}(a,b)) \geq \beta - \delta_{[a,b]}\left(\frac{\lambda r}{b-a}\right) - e^{-k/\sqrt{\epsilon}}.$$

On définit une fonction $\rho^\epsilon_{[a,b]} : \mathbb{R}^+ \to \mathbb{R}$:

$$\rho^\epsilon_{[a,b]}(x) = \frac{1}{\beta - \delta_{[a,b]}(x) - e^{-k/\sqrt{\epsilon}}}$$

Ainsi, nous pouvons déduire qu'il existe $\varepsilon_0 > 0$ et $r_0 > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$ et $\forall r \in (0, r_0]$,

$$P(A_+^c(a,b) \,|\, \widehat{C}_+^{\epsilon,r}(a,b)) \leq e^{-k_1/\sqrt{\epsilon}} + \rho^\epsilon_{[a,b]}\left(\frac{\lambda r}{b-a}\right) k_2\, e^{-k_3 r}$$

□

Remarque 5.14 $\forall x \geq 0$,

$$\rho^\epsilon_{[a,b]}(x) \geq \rho^\epsilon_{[a,b]}\left(\frac{r_{inf}}{b-a}\right) = \frac{1}{\beta - e^{-k/\sqrt{\epsilon}}}.$$

Remarque 5.15 Lorsque $\varepsilon \to 0$, les probabilités d'erreur ne convergent plus vers zéro comme sous l'hypothèse $(HD1)$, mais vers la quantité

$$\rho^{\varepsilon}_{[a,b]}\left(\frac{\lambda r}{b-a}\right) e^{-k_2 r}.$$

Le facteur $\rho^{\varepsilon}_{[a,b]}\left(\frac{\lambda r}{b-a}\right)$ correspond à l'inverse de la probabilité de prendre une décision sur l'intervalle $[a,b]$, il est petit quand le rapport $\frac{r}{b-a}$ est petit. Le second facteur est lié à la probabilité d'erreur du test, il tend vers zéro lorsque r tend vers $+\infty$. Ainsi, ce test ne pourra décider correctement que sur des intervalles de monotonie suffisament longs. Ceci est la différence majeure entre la situation sous $(HD1)$ et sous $(HD2)$.

Procédure du test du rapport de vraisemblance sur les accroissements de l'observation:

Comme dans le test de la variation quadratique, nous introduisons τ_1, le temps à partir duquel nous pouvons prendre une décision pour une probabilité d'erreur donnée.

Soit τ_1, $a < \tau_1 \leq b$. Sur $C^{\varepsilon}(a,b)$: après un certain temps qui peut être aléatoire, $\tau_1 - a$, on décide entre $A_+(a,b)$ et $A_-(a,b)$ d'après le signe de la statistique $L(a,\tau_1)$ calculée à partir des observations recueillies durant l'intervalle de temps $[a,\tau_1]$. Cette procédure se résume de la façon suivante:

On définit deux nouveaux événements tests:

$$\begin{aligned} Q^{\varepsilon,r}_{+}(a,b) &= C^{\varepsilon}(a,b) \cap C^{\varepsilon,r}_{+}(a,\tau_1), \\ Q^{\varepsilon,r}_{-}(a,b) &= C^{\varepsilon}(a,b) \cap C^{\varepsilon,r}_{-}(a,\tau_1). \end{aligned}$$

Corollaire 5.16 *Soit $\lambda > 1$, $r_0 > 0$, il existe $\varepsilon_0 > 0$, $k_1 > 0$, $k_2 > 0$, $k_3 > 0$ et une fonction croissante $\rho^{\varepsilon}_{[a,b]} : \mathbb{R}^+ \to \mathbb{R}$ tels que $\forall r \in (0, r_0]$, $\forall \varepsilon \in (0, \varepsilon_0]$,*

$$\begin{aligned} P(A^c_+(a,b) \,|\, Q^{\varepsilon,r}_{+}(a,b)) &\leq e^{-k_1/\sqrt{\varepsilon}} + \rho^{\varepsilon}_{[a,b]}\left(\frac{\lambda r}{\tau_1 - a}\right) k_2\, e^{-k_3 r} \\ P(A^c_-(a,b) \,|\, Q^{\varepsilon,r}_{-}(a,b)) &\leq e^{-k_1/\sqrt{\varepsilon}} + \rho^{\varepsilon}_{[a,b]}\left(\frac{\lambda r}{\tau_1 - a}\right) k_2\, e^{-k_3 r} \end{aligned}$$

Preuve Démontrons la première inégalité:

$$A_+^c(a,b) \cap Q_+^{\epsilon,r}(a,b) \subset \quad A_+^c(a,\tau_1) \cap C_+^{\epsilon,r}(a,\tau_1) \cup \; A_+(a,\tau_1) \cap A_+^c(a,b) \cap C^\epsilon(a,b).$$

L'inclusion suivante

$$A_+(a,\tau_1) \cap A_+^c(a,b) \cap C^\epsilon(a,b) \subset (\, A_+(a,b) \cup A_-(a,b)\,)^c \cap C^\epsilon(a,b),$$

le théorème 3.4 et la proposition 5.10 impliquent

$$P(A_+^c(a,b) \cap Q_+^{\epsilon,r}(a,b)) \leq e^{-k_1/\sqrt{\epsilon}} + k_2 e^{-k_3 r}, \tag{5.15}$$

avec $k_1 > 0$, $k_2 > 0$, $k_3 > 0$ et ϵ assez petit.

De même que pour le théorème 5.13, on termine la démonstration à l'aide du lemme 5.11.

□

Remarque 5.17 Ce test a été mis en œuvre sous la forme séquentielle dans le chapitre II. Le temps τ_1 est considéré comme un temps d'arrêt. Contrairement à ce que l'étude précédente laisse sous entendre, nous ne cumulerons pas la statistique L^ϵ jusqu'à ce que $R(a,\tau_1) > r$ avec r déterminé pour une probabilité d'erreur donnée, mais nous déterminerons des bornes $l_1 > 0$, $l_2 > 0$, qui permettront de prendre une décision dès que $L^\epsilon(a,\tau_1) \leq -l_1$ ou $L^\epsilon(a,\tau_1) \geq l_2$.

6 Le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy

Dans ce paragraphe, nous présentons un test pour décider sur l'événement C^ϵ, entre "A_+" et "A_-" à partir des sorties $\hat{X}_t^+$ et $\hat{X}_t^-$ des deux filtres de Kalman-Bucy, (FK_+) et (FK_-). Si nous savons que $\langle X_t\,,\,u\rangle$ est différent de zéro sur l'intervalle $[a,b]$, avec une probabilité proche de 1, on peut supposer qu'au bout d'un certain temps $d-a$, avec $a<d<b$, si $\langle X_t\,,\,u\rangle > 0$ (resp.$\langle X_t\,,\,u\rangle < 0$) la loi conditionnelle de X_t sachant $\mathcal{Y}_t$ est proche de la loi gaussienne $N(\hat{X}_t^+, R_\infty^+)$, (resp.$N(\hat{X}_t^-, R_\infty^-)$). On construit une statistique de test $\hat{L}^\epsilon$ du type rapport de vraisemblance, faisant intervenir $\hat{X}_t^+$ et $\hat{X}_t^-$. Sous l'une des hypothèses suivantes:

$$(\overline{HD}1) \qquad |H_+K_\infty^+ - H_-K_\infty^-| \geq \mathcal{O}(1)$$

ou

$$(\overline{HD}2) \qquad \begin{cases} |H_+K_\infty^+ - H_-K_\infty^-| \leq \mathcal{O}(\epsilon) \\ \text{et} \\ \ker\{H_-B_+H_+^{-1} - H_-B_-H_-^{-1}\} \subset \Delta_u. \end{cases}$$

On montre que le signe de cette statistique nous permet de choisir entre "A_-" et "A_+". Evidemment, comme nous l'avons vu précédemment, la qualité du test n'est pas la même dans les situations du type (1) et (2). Sous $(\overline{HD}1)$, la probabilité d'erreur est exponentiellement faible lorsque ϵ est petit. Sous $(\overline{HD}2)$, la probabilité d'erreur dépend de ϵ et de la longueur de l'intervalle de temps $[a,b]$.

Ce test a été proposé dans [6], uniquement dans le cas où l'hypothèse $(HD1)$ est vérifiée. Contrairement à ce qui a été fait dans cet article, nous allons justifier ce test indépendamment des résultats des tests précédents, dans un cadre plus général. La principale difficulté réside dans le fait que le coefficient de diffusion n'est plus supposé constant comme dans [6], mais constant par morceaux.

Etudions tout d'abord le lien entre les hypothèses $(HD1)$-$(HD2)$ et les hypothèses $(\overline{HD}1)$-$(\overline{HD}2)$.

Lemme 6.1

- *L'hypothèse* $(HD1)$ *implique* $(\overline{HD}1)$.

- *Si les matrices H_+G_+ et H_-G_- sont symétriques, l'hypothèse $(HD2)$ implique $(\overline{HD2})$.*

Preuve D'après l'équation (2.11), on note que

$$\begin{aligned} |H_+K_\infty^+ - H_-K_\infty^-| &= |H_+(R_\infty^+H_+^* + G_+) - H_-(R_\infty^-H_-^* + G_-) + \mathcal{O}(\epsilon)| \\ &\geq ||H_+\Sigma_+| - |H_-\Sigma_-|| + \mathcal{O}(\epsilon). \end{aligned}$$

Donc sous $(HD1)$, $|H_+K_\infty^+ - H_-K_\infty^-| \geq \mathcal{O}(1)$.

Sous l'hypothèse $(HD2)$, on a d'après (2.11):

$$\begin{aligned} &H_+(R_\infty^+H_+^* + G_+)(R_\infty^+H_+^* + G_+)^*H_+^* \\ &\qquad = H_-(R_\infty^+H_+^* + G_+)(R_\infty^+H_+^* + G_+)^*H_-^*. \end{aligned} \tag{6.1}$$

La symétrie des matrices H_+G_+ et H_-G_- entraine celle des matrices $H_+(R_\infty^+H_+^* + G_+)$ et $H_-(R_\infty^-H_-^* + G_-)$. L'égalité (6.1) s'écrit

$$[H_+(R_\infty^+H_+^* + G_+)]^2 = [H_-(R_\infty^-H_-^* + G_-)]^2.$$

D'autre part, sachant que ces matrices sont stables (Cf. remarque 2.2), on en déduit qu'elles sont égales et que

$$|H_+K_\infty^+ - H_-K_\infty^-| = \mathcal{O}(\epsilon).$$

□

Remarque 6.2 L'hypothèse de symétrie des matrices H_+G_+ et H_-G_- est évidemment vérifiée dans le cas $n = 1$ et dans le cas d'indépendance des bruits de dynamique et d'observation.

Présentons un cas où l'hypothèse $(\overline{HD1})$ est vérifiée alors que $(HD1)$ ne l'est pas.

Exemple 6.3

$$\begin{array}{llll} H = H_+ = H_- & F_+ = \frac{1}{\sqrt{2}}I_n & G_+ = \frac{1}{\sqrt{2}}I_n \\ \Sigma = \Sigma_+ = \Sigma_- & F_- = I_n & G_- = 0 \end{array}$$

avec la matrice H non symétrique.

On vérifie que

$$H_+\Sigma_+\Sigma_+^*H_+^* = H_-\Sigma_-\Sigma_-^*H_-^*$$

et on pose

$$\begin{aligned} HK_\infty^+ &= HR_\infty^+H^* + \frac{1}{\sqrt{2}}H + o(\varepsilon) \\ HK_- &= HR_\infty^-H^* + o(\varepsilon). \end{aligned}$$

Vu que la matrice $HR_\infty^-H^*$ est symétrique et que H ne l'est pas, on ne peut avoir égalité entre les deux matrices $HR_\infty^+H^* + \frac{1}{\sqrt{2}}H$ et $HR_\infty^-H^*$. Ces matrices ne dépendant pas de ε, on a donc

$$|H_+K_\infty^+ - H_-K_\infty^-| = \mathcal{O}(1).$$

Il est donc important de noter qu'il existe des cas où le test de la variation quadratique ne peut être appliqué mais par le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy nous pouvons décider entre "A_+" et "A_-" avec une probabilité d'erreur exponentiellement faible pour ε petit.

Donnons l'idée générale de ce test. On définit sur $(\Omega, \mathcal{F})$ la loi $\overset{\circ}{P}$ sous laquelle le processus $\{Y_t/\varepsilon\,,\, t>0\}$ est un $\mathcal{F}_t$-processus de Wiener standard:

$$\left.\frac{d\overset{\circ}{P}}{dP}\right|_{\mathcal{F}_t} = \exp\left\{-\frac{1}{\varepsilon}\int_0^t h(X_s)dW_s - \frac{1}{2\varepsilon^2}\int_0^t |h(X_s)|^2\,ds\right\}.$$

Introduisons les notations suivantes:

$$\begin{aligned} \hat{h}_t^+ &= H_+\hat{X}_t^+, \\ \hat{h}_t^- &= H_-\hat{X}_t^-, \\ \hat{h}_t &= E[h(X_t)\,|\,\mathcal{Y}_t]. \end{aligned}$$

Soit $0 < a < d < b$. On introduit les décompositions suivantes, pour $t \in [d, b]$:

$$\begin{aligned} dY_t &= \hat{h}_t^+dt + \varepsilon d\omega_t^+ \\ dY_t &= \hat{h}_t^-dt + \varepsilon d\omega_t^- \end{aligned}$$

avec $\{\omega_t^+ - \omega_d^+ ; d \le t \le b\}$ (resp. $\{\omega_t^- - \omega_d^- ; d \le t \le b\}$) un processus de Wiener standard sous la loi $\overset{\circ}{P}{}^+$ (resp. $\overset{\circ}{P}{}^-$). On définit les probabilités $\overset{\circ}{P}{}^+$ et $\overset{\circ}{P}{}^-$ sur $(\Omega, \mathcal{F}_b^d)$ en utilisant le lemme 2.8 et le théorème de Girsanov:

$$\frac{d\overset{\circ}{P}{}^+}{d\overset{\circ}{P}} = \exp\left\{\frac{1}{\varepsilon^2}\int_d^b \hat{h}_s^+ dY_s - \frac{1}{2\varepsilon^2}\int_d^b |\hat{h}_s^+|^2 ds\right\}$$

$$\frac{d\overset{\circ}{P}{}^-}{d\overset{\circ}{P}} = \exp\left\{\frac{1}{\varepsilon^2}\int_d^b \hat{h}_s^- dY_s - \frac{1}{2\varepsilon^2}\int_d^b |\hat{h}_s^-|^2 ds\right\}$$

Ainsi, nous obtenons la statistique :

$$\hat{L}^\varepsilon = \ln \frac{d\overset{\circ}{P}{}^+}{d\overset{\circ}{P}{}^-} = \frac{1}{\varepsilon^2}\int_d^b (\hat{h}_t^+ - \hat{h}_t^-) dY_t - \frac{1}{2\varepsilon^2}\int_d^b (|\hat{h}_t^+|^2 - |\hat{h}_t^-|^2)\, dt. \qquad (6.2)$$

On pose

$$\begin{aligned} X_t^1 &= \langle X_t , u\rangle, \\ X_t^{1,+} &= \sup(0, X_t^1), \\ X_t^{1,-} &= \sup(0, -X_t^1). \end{aligned}$$

Réécrivons le système (2.3):

$$\begin{cases} dX_t &= [1_{\{X_t^1 \ge 0\}} B_+ X_t + 1_{\{X_t^1 < 0\}} B_- X_t]\, dt \\ &\quad + 1_{\{X_t^1 \ge 0\}}(F_+ dV_t + G_+ dW_t) + 1_{\{X_t^1 < 0\}}(F_- dV_t + G_- dW_t), \\ dY_t &= [1_{\{X_t^1 \ge 0\}} H_+ X_t + 1_{\{X_t^1 < 0\}} H_- X_t]\, dt + \varepsilon\, dW_t. \end{cases} \qquad (6.3)$$

Nous nous inspirons de cette décomposition pour introduire deux processus intermédiaires $\{M_t^+ ; t \ge 0\}$ et $\{M_t^- ; t \ge 0\}$. Par un changement de probabilité, nous construisons un problème de filtrage linéaire où le processus $\{M_t^+ ; t \ge 0\}$, respectivement $\{M_t^- ; t \ge 0\}$, est le processus à estimer et $\{Y_t ; t \ge 0\}$ le processus observé.

Intéressons nous au cas positif. D'après la formule d'Itô-Tanaka, on a:

$$\begin{aligned} dX_t^{1,+} &= 1_{\{X_t^{1,+} \ne 0\}} \langle B_+ X_t , u\rangle\, dt \\ &\quad + 1_{\{X_t^{1,+} \ne 0\}} \langle F_+ dV_t + G_+ dW_t , u\rangle + \frac{1}{2} dL_t^0, \end{aligned} \qquad (6.4)$$

avec $\{L_t^0\,;\,t\geq 0\}$ le processus temps local en zéro de la semi-martingale réelle $\{X_t^1\,;t\geq 0\}$. L'égalité (6.4) se réécrit:

$$\begin{aligned} dX_t^{1,+} = \quad & \mathbf{1}_{\{X_t^{1,+}\neq 0\}}\langle B_+X_t\,,\,u\rangle\,dt + \langle F_+dV_t + G_+dW_t\,,\,u\rangle \\ - \quad & \mathbf{1}_{\{X_t^{1,+}=0\}}\langle F_+dV_t + G_+dW_t\,,\,u\rangle + \frac{1}{2}dL_t^0. \end{aligned}$$

Soit $0 < a < d < b$, on introduit un processus $\{X_t^+\,;\,t\geq 0\}$ vérifiant $\langle X_t^+\,,\,u\rangle = X_t^{1,+}\,,\,\forall t\geq a$:

$$\begin{cases} dX_t^+ &= \mathbf{1}_{\{X_t^{1,+}\neq 0\}}B_+X_tdt + F_+dV_t + G_+dW_t \\ & \quad -\mathbf{1}_{\{X_t^{1,+}=0\}}[F_+dV_t + G_+dW_t] + \frac{1}{2}dL_t^0\,u, \quad t\geq a, \\ X_t^+ &= X_t, \quad t\leq a. \end{cases}$$

En remarquant que

$$B_+X_t^+ - \mathbf{1}_{\{X_t^{1,+}\neq 0\}}B_+X_t = B_+Q(X_t^+ - \mathbf{1}_{\{X_t^{1,+}\neq 0\}}X_t),$$

on obtient:

$$\begin{aligned} dX_t^+ &= B_+X_t^+dt - B_+Q(X_t^+ - \mathbf{1}_{\{X_t^{1,+}\neq 0\}}X_t)dt \\ &\quad + F_+dV_t + G_+dW_t \\ &\quad -\mathbf{1}_{\{X_t^{1,+}=0\}}[F_+dV_t + G_+dW_t] + \frac{1}{2}dL_t^0\,u, \quad t\geq a. \end{aligned}$$

On décompose X_t^+ de la façon suivante:

$$X_t^+ = M_t^+ - \psi_t^a - \varphi_t^a + \phi_t^a,$$

où

- $\{M_t^+\,;\,t\geq 0\}$ est le processus défini par

$$\begin{cases} dM_t^+ &= B_+M_t^+dt + B_+(-\psi_t^a - \varphi_t^a + \phi_t^a)dt + F_+dV_t + G_+dW_t, \quad t\geq a, \\ M_t^+ &= X_t, \quad t\leq a. \end{cases} \tag{6.5}$$

- $\psi_t^a = \int_a^{t\vee a} B_+Q(X_s^+ - \mathbf{1}_{\{X_s^{1,+}\neq 0\}}X_s)\,ds$

- $\varphi_t^a = F_+ \int_a^{t\vee a} 1_{\{X_s^{1,+}=0\}} dV_s + G_+ \int_a^{t\vee a} 1_{\{X_s^{1,+}=0\}} dW_s.$
- $\phi_t^a = \frac{1}{2}(L^0_{t\vee a} - L^0_a)u.$

Considérons le système suivant, $\forall t \geq a$:

$$\begin{cases} dM_t^+ &= \left[B_+ M_t^+ + B_+(-\psi_t^a - \varphi_t^a + \phi_t^a\right] dt + F_+ dV_t + G_+ dW_t, \\ dY_t &= \left[H_+ M_t^+ + H_+(-\psi_t^a - \varphi_t^a + \phi_t^a) + 1_{\{X_t^1 \leq 0\}} H_- X_t\right] dt + \varepsilon\, dW_t. \end{cases} \tag{6.6}$$

On pose

$$\alpha_t^\varepsilon = F_+^{-1}\left(B_+(\psi_t^a + \varphi_t^a - \phi_t^a) - \frac{1}{\varepsilon} G_+ \beta_t\right), \tag{6.7}$$

$$\beta_t = H_+(\psi_t^a + \varphi_t^a - \phi_t^a) - 1_{\{X_t^1 \leq 0\}} H_- X_t, \tag{6.8}$$

et on définit les processus $\{\tilde{V}_t^+, t \geq 0\}$ et $\{\tilde{W}_t^+, t \geq 0\}$ de la manière suivante:

$$\tilde{V}_t^+ = V_t - \int_a^{t\vee a} \alpha_s^\varepsilon ds,$$

$$\tilde{W}_t^+ = W_t - \frac{1}{\varepsilon}\int_a^{t\vee a} \beta_s ds.$$

Soit

$$Z_t = \exp\left\{\int_a^{t\vee a} \alpha_s^\varepsilon dV_s + \frac{1}{\varepsilon}\int_a^{t\vee a} \beta_s dW_s - \frac{1}{2}\int_a^{t\vee a} (\alpha_s^\varepsilon)^2 ds - \frac{1}{2\varepsilon^2}\int_a^{t\vee a} \beta_s^2 ds\right\},$$

Le lemme 2.7, la majoration (3.9) du moment exponentiel du processus temps local sur $[0,T]$ et l'inégalité de Cauchy-Schwartz impliquent

$$\mathbf{E}\left(\exp c(|\alpha_t^\varepsilon|^2 + |\beta_t/\varepsilon|^2)\right) < +\infty, \quad \forall t \in [0,T], \text{ pour un certain } c > 0,$$

donc $\mathbf{E}(Z_t) = 1$, $t \geq 0$. On peut alors définir la loi de probabilité $\tilde{P}^+$ sur $(\Omega, \mathcal{F}_b)$ par

$$\frac{d\tilde{P}^+}{dP} = Z_b$$

telle que $\{(\tilde{V}_t^+, \tilde{W}_t^+);\, 0 \leq t \leq b\}$ est un $\tilde{P}^+$-processus de Wiener standard à valeurs dans $\mathbb{R}^{2n}$. On note que $\tilde{P}^+$ dépend de ε.

Remarque 6.4 Les restrictions de P et $\tilde{P}^+$ à $(\Omega, \mathcal{F}_a)$ coincident, et

$$\forall D \subset A_+(a,b), \quad P(D) = \tilde{P}^+(D).$$

Pour $t \geq a$, le système (6.6) s'écrit:

$$\begin{cases} dM_t^+ &= B_+ M_t^+ dt + F_+ d\tilde{V}_t^+ + G_+ d\tilde{W}_t^+, \\ dY_t &= H_+ M_t^+ dt + \varepsilon\, d\tilde{W}_t^+. \end{cases}$$

On introduit la décomposition suivante:

$$dY_t = H_+ \tilde{M}_t^+ dt + \varepsilon\, d\tilde{\omega}_t^+, \tag{6.9}$$

où $\tilde{M}_t^+ = \tilde{\mathbf{E}}^+(M_t^+ \mid \mathcal{Y}_t)$ et $\tilde{\omega}_t^+ = \tilde{W}_t^+ + \int_0^t H_+(M_s^+ - \tilde{M}_s^+)ds$. Le processus $\{\tilde{\omega}_t^+ \,;\, 0 \leq t \leq b\}$ est un $\mathcal{F}_t$-processus de Wiener standard, dit processus d'innovation sous la loi $\tilde{P}^+$. En utilisant (6.9), on donne une autre expression de la statistique $\hat{L}^\varepsilon$ définie en (6.2) :

$$\begin{aligned} \hat{L}^\varepsilon &= \frac{1}{2}\langle N^+\rangle_b + N_b^+ \\ &\quad + \frac{1}{\varepsilon^2}\int_a^b (\hat{h}_t^+ - \hat{h}_t^-)(H_+ \tilde{M}_t^+ - H_+ \hat{X}_t^+)\, dt, \end{aligned} \tag{6.10}$$

avec

$$N_t^+ = \frac{1}{\varepsilon}\int_a^t (\hat{h}_s^+ - \hat{h}_s^-) d\tilde{\omega}_s^+.$$

De la même façon que nous avons défini la probabilité $\hat{P}^+$, on définit la probabilité $\tilde{P}^-$ sur $(\Omega, \mathcal{F}_b)$. La statistique de test s'écrit :

$$\begin{aligned} \hat{L}^\varepsilon &= -\frac{1}{2}\langle N^-\rangle_b + N_b^- \\ &\quad - \frac{1}{\varepsilon^2}\int_a^b (\hat{h}_t^+ - \hat{h}_t^-)(H_- \tilde{M}_t^- - H_- \hat{X}_t^-)\, dt, \end{aligned} \tag{6.11}$$

avec

$$N_t^- = \frac{1}{\varepsilon}\int_a^t (\hat{h}_s^+ - \hat{h}_s^-) d\tilde{\omega}_s^-.$$

et $\{\tilde{\omega}_t^- \,,\, 0 \leq t \leq b\}$ est un processus de Wiener standard sous la loi $\tilde{P}^-$.

Remarque 6.5 $\langle N^+\rangle_b = \langle N^-\rangle_b$.

Montrons que sur l'événement $C^\epsilon(a,b)$, la statistique de test $\hat{L}^\epsilon$ permet de choisir entre les deux alternatives "$A_+(a,b)$" et "$A_-(a,b)$". Nous considérerons le "cas positif", le "cas négatif" étant similaire.

Premièrement, on s'intéresse à la différence $|\tilde{M}_t^+ - \hat{X}_t^+|$ sur l'intervalle de temps $[e,b]$, avec $0 < a < e < d < b$. On remarque que sur $A_+(a,b)$, les processus $\{\hat{X}_t^+ ; t \geq e\}$ et $\{\tilde{M}_t^+ ; t \geq e\}$ sont solutions du même problème de filtrage linéaire avec des conditions initiales différentes. La loi initiale du second n'est pas gaussienne, mais pour ϵ petit, le filtre étant à mémoire courte, il est pratiquement insensible à la loi initiale à partir d'un instant $e > a$.

Lemme 6.6 *Soit* $0 < a < e < d < b$,
$\forall \theta > 0$, $\mu > 0$, $q \geq 1$, *il existe* $k_{q,\mu} > 0$, $\varepsilon_{q,\mu} > 0$ *tels que* $\forall \epsilon \in (0, \varepsilon_{q,\mu}]$,

$$P\left(\left\{\int_e^b |\tilde{M}_t^+ - \hat{X}_t^+|^q \, dt \geq \theta \epsilon^\mu\right\} \cap A_+(a,b)\right) \leq e^{-k_{q,\mu}/\epsilon}.$$

Preuve Soit $\mathcal{Y}_t^a = \sigma(Y_s - Y_a\,,\, a \leq s \leq t)$, $t \geq a$, premièrement nous montrons que sous la loi $\tilde{P}^+$, les tribus $\mathcal{Y}_a$ et $\sigma(M_t^+) \vee \mathcal{Y}_t^a$ sont conditionnellement indépendantes sachant $\sigma(M_a^+)$ pour $t \geq a$. On rappelle que $M_a^+ = X_a$. Il est clair qu'il est suffisant de montrer que

$$\sigma(\, X_0, V_s, W_s,\, s \leq a\,) \perp\!\!\!\perp_{\sigma(X_a)} \sigma(\, X_a, V_s - V_a, W_s - W_a,\, s \geq a\,).$$

En utilisant l'implication suivante:

$$\text{soit les tribus } \mathcal{F},\ \mathcal{G},\ \mathcal{H}, \qquad \mathcal{H} \perp\!\!\!\perp \mathcal{F} \vee \mathcal{G} \Longrightarrow \mathcal{F} \perp\!\!\!\perp_{\mathcal{G}} \mathcal{G} \vee \mathcal{H},$$

et le fait que

$$\sigma(\, V_s - V_a, W_s - W_a,\, s \geq a\,) \perp\!\!\!\perp \sigma(\, X_0, X_a, V_s, W_s,\, s \leq a\,),$$

on obtient l'indépendance conditionnelle demandée. Nous en déduisons les égalités suivantes:

$$\begin{aligned} \tilde{M}_t^+ &= \tilde{\mathbf{E}}^+(M_t^+ \mid \mathcal{Y}_t) \\ &= \tilde{\mathbf{E}}^+\left(\tilde{\mathbf{E}}^+(M_t^+ \mid \mathcal{Y}_t, X_a) \mid \mathcal{Y}_t\right) \\ &= \tilde{\mathbf{E}}^+\left(\tilde{\mathbf{E}}^+(M_t^+ \mid \mathcal{Y}_t^a, X_a) \mid \mathcal{Y}_t\right). \end{aligned} \tag{6.12}$$

Soit $\tilde{M}^+_{a,t} = \tilde{\mathbf{E}}^+(M^+_t \mid \mathcal{Y}^a_t, X_a)$, $\tilde{M}^+_{a,t}$ est la sortie du filtre de Kalman suivant:

$$\begin{cases} d\tilde{M}^+_{a,t} &= B_+\tilde{M}^+_{a,t}dt + \frac{1}{\epsilon}(G_+ + \frac{1}{\epsilon}R^+_{a,t}H^*_+)(dY_t - H_+\tilde{M}^+_{a,t}dt), \\ \tilde{M}^+_{a,a} &= X_a, \\ \frac{dR^+_{a,t}}{dt} &= F_+F^*_+ + (B_+ - \frac{1}{\epsilon}G_+H_+)R^+_{a,t} + R^+_{a,t}(B_+ - \frac{1}{\epsilon}G_+H_+)^* \\ &\qquad -\frac{1}{\epsilon^2}R^+_{a,t}H^*_+H_+R^+_{a,t}, \\ R^+_{a,a} &= 0. \end{cases} \tag{6.13}$$

On pose $K^+_{a,t} = G_+ + \frac{1}{\epsilon}R^+_{a,t}H^*_+$. D'après (2.7), il existe $k > 0$ tel que

$$|R^+_{a,t} - R^+_\infty| \leq e^{-k/\epsilon}, \text{ pour } t \in [e, b]. \tag{6.14}$$

Les égalités (6.13) et (2.12) impliquent:

$$\begin{cases} d(\hat{X}^+_t - \tilde{M}^+_{a,t}) &= (B_+ - \frac{1}{\epsilon}K^+_\infty H_+)(\hat{X}^+_t - \tilde{M}^+_{a,t})dt + (K^+_\infty - K^+_{a,t})d\tilde{\omega}^+_t, \\ \hat{X}^+_a - \tilde{M}^+_{a,a} &= \hat{X}^+_a - X_a. \end{cases} \tag{6.15}$$

Vu que la matrice $\Lambda = \epsilon B_+ - K^+_\infty H_+$ est stable, en utilisant (6.14) et (6.15), on montre que pour $q \geq 1$, $t \geq e$, il existe $k_q > 0$, $\nu_q > 0$, $\varepsilon_q > 0$ tels que $\forall \epsilon \in (0, \varepsilon_0]$,

$$\tilde{\mathbf{E}}^+|\hat{X}^+_t - \tilde{M}^+_{a,t}|^q \leq \nu_q(1 + \tilde{\mathbf{E}}^+|X_a|^q + \tilde{\mathbf{E}}^+|\hat{X}^+_a|^q)e^{-k_q/\epsilon}.$$

Sous les lois P et $\tilde{P}^+$, les vecteurs X_a et $\hat{X}^+_a$ ont tous leurs moments finis d'après le lemme 2.7,

$$\tilde{\mathbf{E}}^+|X_a|^q = E|X_a|^q < +\infty,$$
$$\tilde{\mathbf{E}}^+|\hat{X}^+_a|^q = E|\hat{X}^+_a|^q < +\infty.$$

On en déduit

$$\tilde{\mathbf{E}}^+|\hat{X}^+_t - \tilde{M}^+_{a,t}|^q \leq \alpha_q e^{-k_q/\epsilon},$$

pour $t \geq e$ et un certain $\alpha_q > 0$. D'après (6.12) et l'inégalité de Jensen, on a, pour $t \geq e$:

$$\begin{aligned} \tilde{\mathbf{E}}^+|\hat{X}^+_t - \tilde{M}^+_t|^q &= \tilde{\mathbf{E}}^+\left((\tilde{\mathbf{E}}^+(|\hat{X}^+_t - \tilde{M}^+_{a,t}| \mid \mathcal{Y}_t))^q\right) \\ &\leq \tilde{\mathbf{E}}^+|\hat{X}^+_t - \tilde{M}^+_{a,t}|^q \\ &\leq \alpha_q e^{-k_q/\epsilon}. \end{aligned} \tag{6.16}$$

Par le théorème de Fubini, on obtient

$$\tilde{\mathbf{E}}^+ \int_e^b |\hat{X}_t^+ - \tilde{M}_{a,t}^+|^q dt \leq \alpha_q (b-a) e^{-k_q/\epsilon}. \tag{6.17}$$

L'inégalité de Markov et la remarque 6.4 nous permettent de conclure.

□

Intéressons nous au terme de précision.

Lemme 6.7 *Soit* $0 < a < e < d < b$,

Sous $(\overline{HD}1)$: *il existe* $r_{inf}^1 < 0$, $\varepsilon_0 > 0$, $k > 0$ *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\frac{1}{\varepsilon}\int_d^b \left|\hat{h}_s^+ - \hat{h}_s^-\right|^2 ds < r_{inf}^1\right\} \cap C^\epsilon(a,b)\right) \leq e^{-k/\sqrt{\epsilon}}.$$

Sous $(\overline{HD}2)$: *il existe* $r_{inf}^2 > 0$, $\varepsilon_0 > 0$, $k > 0$ *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\frac{1}{\varepsilon^2}\int_d^b \left|\hat{h}_s^+ - \hat{h}_s^-\right|^2 ds < r_{inf}^2\right\} \cap C^\epsilon(a,b)\right) \leq e^{-k/\sqrt{\epsilon}}.$$

Preuve On pose $Z_t = \hat{h}_t^+ - \hat{h}_t^-$. D'après le théorème 3.4, il suffit de considérer les termes de précision sur l'événement $C^\epsilon(a,b) \cap (A_-(a,b) \cup A_+(a,b))$. Nous raisonnons sur $C^\epsilon(a,b) \cap A_+(a,b)$, la démonstration étant similaire sur $C^\epsilon(a,b) \cap A_-(a,b)$.

$$\begin{aligned} dH_+\hat{X}_t^+ &= H_+B_+\hat{X}_t^+dt + \frac{1}{\varepsilon}H_+K_\infty^+(dY_t - H_+\hat{X}_t^+dt), \\ dH_-\hat{X}_t^- &= H_-B_-\hat{X}_t^-dt + \frac{1}{\varepsilon}H_-K_\infty^-(dY_t - H_-\hat{X}_t^-dt). \end{aligned}$$

En utilisant la décomposition suivante

$$dY_t = H_+\tilde{M}_t^+dt + \varepsilon d\tilde{\omega}_t^+, \quad t \geq a. \tag{6.18}$$

on obtient

$$\begin{aligned} dZ_t &= \frac{1}{\varepsilon}\Lambda Z_t dt + (H_+B_+H_+^{-1} - H_-B_-H_-^{-1})H_+\hat{X}_t^+dt \\ &\quad + \frac{1}{\varepsilon}(H_+K_\infty^+ - H_-K_\infty^-)(H_+\tilde{M}_t^+ - H_+\hat{X}_t^+)dt \\ &\quad + (H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_t^+, \end{aligned} \tag{6.19}$$

avec $\Lambda = \varepsilon H_-B_-H_-^{-1} - H_-K_\infty^-$, matrice stable pour ε assez petit. (Voir la remarque 2.2).

Sous $(\overline{HD}1)$: On décompose Z_t de la manière suivante:

$$Z_t = Z_t^{(1)} + Z_t^{(2)},$$

avec

$$\begin{cases} dZ_t^{(1)} &= \frac{1}{\epsilon}\Lambda Z_t^{(1)}dt + (H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_t^+, \\ Z_e^{(1)} &= 0. \end{cases}$$

$$\begin{cases} dZ_t^{(2)} = & \frac{1}{\epsilon}\Lambda Z_t^{(2)}dt + (H_+B_+H_+^{-1} - H_-B_-H_-^{-1})H_+\hat{X}_t^+\,dt \\ & + \frac{1}{\epsilon}(H_+K_\infty^+ - H_-K_\infty^-)(H_+\tilde{M}_t^+ - H_+\hat{X}_t^+)\,dt \\ Z_e^{(2)} = & Z_e. \end{cases}$$

Montrons que $Z_t^{(2)}$ est de faible contribution:

$$\begin{aligned} \left|Z_t^{(2)}\right| \;\leq\; & \left|e^{\frac{1}{\epsilon}\Lambda(t-e)}\right||Z_e| + \left|\int_e^b e^{\frac{1}{\epsilon}\Lambda(t-s)}(H_+B_+H_+^{-1} - H_-B_-H_-^{-1})H_+\hat{X}_s^+ds\right| \\ & + \frac{1}{\varepsilon}\left|\int_e^b e^{\frac{1}{\epsilon}\Lambda(t-s)}(H_+K_\infty^+ - H_-K_\infty^-)(H_+\tilde{M}_s^+ - H_+\hat{X}_s^+)ds\right|. \end{aligned}$$

La stabilité de la matrice Λ entraine que le 1^{ier} terme est exponentiellement petit pour $t \geq d$, et du lemme 2.8, on déduit que le second est au plus d'ordre $\mathcal{O}(\varepsilon)$. Intéressons nous au dernier terme. On note

$$\psi_t = \frac{1}{\varepsilon}\int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)}(H_+K_\infty^+ - H_-K_\infty^-)(H_+\tilde{M}_s^+ - H_+\hat{X}_s^+)ds.$$

Les inégalités de Cauchy-Schwartz et de Jensen impliquent:

$$\begin{aligned} |\psi_t|^2 \;\leq\; & \frac{1}{\varepsilon^2}|H_+K_\infty^+ - H_-K_\infty^-|^2|H_+|^2\int_e^t \left|e^{\frac{1}{\epsilon}\Lambda(t-s)}\right|^2 ds \int_e^t |\tilde{M}_s^+ - \hat{X}_s^+|^2 ds \\ \leq\; & \frac{k}{\varepsilon}\int_e^t |\tilde{M}_s^+ - \hat{X}_s^+|^2 ds. \end{aligned}$$

Du lemme 6.6, on conclut que $\forall\theta > 0$, il existe $k > 0$, $\varepsilon_0 > 0$ tels que $\forall\varepsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\frac{1}{\varepsilon}\sup_{[d,b]}|\psi_t|^2\,ds \geq \theta\right\} \cap A_+(a,b)\right) \leq e^{-k/\epsilon}.$$

Considérons $Z_t^{(1)}$:

$$Z_t^{(1)} = \int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)}(H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_s^+.$$

Pour compléter la démonstration sous l'hypothèse $(\overline{HD}1)$, montrons qu'il existe $r_1 > 0$, $k = k(r_1) > 0$, $\varepsilon_0 = \varepsilon_0(r_1) > 0$ tels que $\forall \epsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\frac{1}{\varepsilon}\int_d^b \left|Z_t^{(1)}\right|^2 dt < r_1\right\} \cap A_+(a,b)\right) \leq e^{-k/\sqrt{\epsilon}}. \tag{6.20}$$

Le processus $\{Z_t^{(1)}, t \geq e\}$ vérifie l'équation différentielle stochastique suivante:

$$\begin{cases} dZ_t^{(1)} &= \frac{1}{\epsilon}\Lambda Z_t^{(1)}dt + (H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_t^+, \ t \geq e; \\ Z_e^{(1)} &= 0. \end{cases}$$

La matrice Λ étant stable, il existe une solution unique Π, $\Pi = \Pi^* > 0$, de l'équation matricielle:

$$\Lambda^*\Pi + \Pi\Lambda = -I_n. \tag{6.21}$$

D'après la formule d'Itô et l'équation (6.21), on a:

$$\begin{aligned} \frac{1}{\varepsilon}\int_d^b \left|Z_t^{(1)}\right|^2 dt &= -Z_b^{(1),*}\Pi Z_b^{(1)} + Z_d^{(1),*}\Pi Z_d^{(1)} + N_b - N_d \\ &\quad + (b-d)\left|\Pi^{\frac{1}{2}}(H_+K_\infty^+ - H_-K_\infty^-)\right|^2 \end{aligned}$$

avec

$$N_t = 2\int_e^t Z_s^{(1),*}\Pi(H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_s^+.$$

On remarque que sous $(\overline{HD}1)$, il existe $\tilde{r}_1 > 0$ tel que

$$(b-d)\left|\Pi^{\frac{1}{2}}(H_+K_\infty^+ - H_-K_\infty^-)\right|^2 \geq \tilde{r}_1.$$

Prennons $r_1 = \frac{\tilde{r}_1}{2}$. Pour obtenir (6.20), établissons que $\forall \theta > 0$, on a:

$$P\left(\left\{\left|-Z_b^{(1),*}\Pi Z_b^{(1)} + Z_d^{(1),*}\Pi Z_d^{(1)} + N_b - N_d\right| > \theta\right\} \cap A_+(a,b)\right) \leq e^{-k/\sqrt{\epsilon}}.$$

$$
\begin{aligned}
& P\left(\left\{\left|-Z_b^{(1),*}\Pi Z_b^{(1)}+Z_d^{(1),*}\Pi Z_d^{(1)}+N_b-N_d\right|>\theta\right\}\cap A_+(a,b)\right)\\
\leq\quad & P\left(\left\{\sup_{[d,b]}|Z_t^{(1)}|^2>\frac{\theta}{4|\Pi|}\right\}\cap A_+(a,b)\right)\\
+\quad & P\left(\left\{|N_b-N_d|>\frac{\theta}{2}\right\}\cap A_+(a,b)\right)
\end{aligned}
$$

De la même façon qui nous avons obtenu la majoration (3.8) dans la démonstration de la proposition 3.1, on montre que

$$
P\left(\left\{\sup_{[d,b]}|Z_t^{(1)}|^2>\frac{\theta}{4|\Pi|}\right\}\cap A_+(a,b)\right)\leq e^{-\frac{k\theta}{\varepsilon}}, \tag{6.22}
$$

pour un certain $k>0$ et ε suffisament petit.

Intéressons nous à la probabilité

$$
P(\{|N_b-N_d|>\frac{\theta}{2}\}\cap A_+(a,b)).
$$

Soit l'événement $D=\{\langle N\rangle_b-\langle N\rangle_d>\frac{\theta\sqrt{\varepsilon}}{2}\}$. D'après (6.22),

$$
P(D\cap A_+(a,b))\leq e^{-k/\sqrt{\varepsilon}}.
$$

On considère donc la probabilité de l'événement

$$
\{|N_b-N_d|>\frac{\theta}{2}\}\cap A_+(a,b)\cap D^c.
$$

On introduit la $\tilde{P}^+$-martingale exponentielle

$$
H_t=\exp\left\{\frac{1}{\sqrt{\varepsilon}}(N_t-N_d)-\frac{1}{2\varepsilon}(\langle N\rangle_t-\langle N\rangle_d)\right\}.
$$

L'inégalité de Doob et le fait que $\tilde{\mathrm{E}}^+(H_b)=1$ impliquent

$$
\begin{aligned}
& P\left(\left\{N_b-N_d>\frac{\theta}{2}\right\}\cap A_+(a,b)\cap D^c\right)\\
\leq\ & \tilde{P}^+\left(\left\{N_b-N_d>\frac{\theta}{2}\right\}\cap D^c\right)\\
\leq\ & \tilde{P}^+\left(\left\{H_b>e^{\theta/4\sqrt{\varepsilon}}\right\}\right)\ \leq\ e^{-\theta/4\sqrt{\varepsilon}}.
\end{aligned}
$$

De même, on démontre que

$$P\left(\left\{-N_b + N_d > \frac{\theta}{2}\right\} \cap A_+(a,b) \cap D^c\right) \leq e^{-\theta/4\sqrt{\epsilon}}.$$

Ainsi dans la situation où l'hypothèse $(\overline{HD}1)$ est satisfaite, nous pouvons conclure.

Sous $(\overline{HD}2)$: On décompose Z_t de la manière suivante:

$$Z_t = Z_t^{(1)} + Z_t^{(2)},$$

avec

$$\begin{cases} dZ_t^{(1)} &= \frac{1}{\epsilon}\Lambda Z_t^{(1)}dt + (H_+B_+H_+^{-1} - H_-B_-H_-^{-1})H_+\hat{X}_t^+dt, \\ Z_e^{(1)} &= 0, \end{cases}$$

$$\begin{cases} dZ_t^{(2)} &= \frac{1}{\epsilon}\Lambda Z_t^{(2)}dt + \frac{1}{\epsilon}(H_+K_\infty^+ - H_-K_\infty^-)(H_+\tilde{M}_t^+ - H_+\hat{X}_t^+)dt \\ & \qquad\qquad +(H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_t^+, \\ Z_e^{(2)} &= Z_e. \end{cases}$$

Montrons que $Z_t^{(2)}$ est de faible contribution:

$$\begin{aligned} \left|Z_t^{(2)}\right| \leq & \left|e^{\frac{1}{\epsilon}\Lambda(t-e)}\right||Z_e| \\ & + \left|\frac{1}{\epsilon}\int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)}(H_+K_\infty^+ - H_-K_\infty^-)H_+(\tilde{M}_s^+ - \hat{X}_s^+)ds\right| \\ & + \left|\int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)}(H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_s^+\right|. \end{aligned}$$

La stabilité de la matrice Λ et le lemme 6.6 nous permettent de négliger les deux premiers termes de droite de l'inégalité. Intéressons nous au troisième.

On pose

$$\psi_t = \int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)}(H_+K_\infty^+ - H_-K_\infty^-)d\tilde{\omega}_s^+,$$

où $H_+K_\infty^+ - H_-K_\infty^-$ est d'ordre $\mathcal{O}(\epsilon)$ sous $(\overline{HD}2)$. A l'aide du raisonnement qui nous a permis d'établir l'inégalité (3.8) et la remarque 6.4, on montre que

$$P\left(\left\{\frac{1}{\epsilon^2}\sup_{[0,b]}|\psi_t|^2 > \theta\right\} \cap A_+(a,b)\right) \leq e^{-k\theta/\epsilon},$$

pour ϵ suffisament petit.

Considérons $Z_t^{(1)}$. On pose $D = H_+ B_+ H_+^{-1} - H_- B_- H_-^{-1}$.

$$Z_t^{(1)} = \int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)} D H_+ \hat{X}_s^+ ds.$$

En intégrant par parties et en utilisant la décomposition (6.18), on obtient

$$\begin{aligned}
& \int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)} D H_+ \hat{X}_s^+ ds \\
= \quad & -\epsilon\Lambda^{-1} D H_+ \hat{X}_t^+ + \epsilon\Lambda^{-1} e^{\frac{1}{\epsilon}\Lambda(t-e)} D H_+ \hat{X}_e^+ \\
+ \quad & \epsilon\Lambda^{-1} \int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)} D H_+ B_+ \hat{X}_s^+ ds \\
+ \quad & \epsilon\Lambda^{-1} \int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)} D H_+ K_\infty^+ H_+ (\tilde{M}_s^+ - \hat{X}_s^+) ds \\
+ \quad & \epsilon\Lambda^{-1} \int_e^t e^{\frac{1}{\epsilon}\Lambda(t-s)} D H_+ K_\infty^+ d\tilde{\omega}_s^+. \qquad (6.23)
\end{aligned}$$

Si on note T_t les quatre derniers termes de (6.23), on montre par des raisonnements similaires à ceux qui précèdent, que $\forall \theta > 0$, $\exists k > 0$, $\varepsilon_0 > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$

$$P\left(\left\{\frac{1}{\varepsilon^2} \sup_{[d,b]} |T_t|^2 > \theta\right\} \cap A_+(a,b)\right) \leq e^{-k/\sqrt{\epsilon}}.$$

Le premier terme est donc prépondérant. Pour conclure sous $(\overline{HD}2)$, montrons que sur $C^\epsilon(a,b)$, il existe $r_2 > 0$ tel que

$$\frac{1}{\varepsilon^2} \int_d^b \left|\varepsilon\Lambda^{-1} D H_+ \hat{X}_t^+\right|^2 dt > r_2 \qquad (6.24)$$

Sur $C^\epsilon(a,b)$, on sait que $\left|\langle \hat{X}_t^+ , u\rangle\right| > c$, $\forall t \in [a,b]$. Donc, d'après l'hypothèse $(\overline{HD}2)$ et le fait que $h(\Delta_u) = \Delta_u$, il existe $c_{inf} > 0$ tel que

$$\inf_{[a,b]} \left|\Lambda^{-1} D H_+ \hat{X}_t^+\right|^2 > c_{inf}. \qquad (6.25)$$

D'où, si nous prennons $r_2 \leq (b-d)\, c_{inf}$, la minoration (6.24) est vérifiée.

□

Remarque 6.8 La différence cruciale entre les deux situations apparaît dans ce lemme. Sur C^ϵ, avec une probabilité proche de 1, le terme de précision $\frac{1}{\epsilon^2}\int_d^b |\hat{h}_s^+ - \hat{h}_s^-|^2 ds$ est minoré par:

- r_{inf}^1/ϵ sous $(\overline{HD}1)$
- r_{inf}^2 sous $(\overline{HD}2)$.

Remarque 6.9 $r_{inf}^2 < r_2 \leq (b-d)\, c_{inf}$ avec c_{inf} défini en (6.25).

Sous l'hypothèse $(\overline{HD}2)$, on définit une fonction croissante $\hat{\delta}_{[a,b]}$:

$$\hat{\delta}_{[a,b]} : \mathbb{R}^+ \longrightarrow [0,1[, \tag{6.26}$$

$$\forall x \geq 0, \qquad \hat{\delta}_{[a,b]}(x) = P\left(\left\{\inf_{[a,b]} \left|\Lambda^{-1} D H_+ \hat{X}_t^+\right|^2 \leq x\right\} \cap C^\epsilon(a,b)\right). \tag{6.27}$$

D'après (6.25), il existe $c_{inf} > 0$ tel que

$$\forall x \in [0, c_{inf}], \qquad \hat{\delta}_{[a,b]}(x) = 0.$$

On montre le résultat suivant:

Lemme 6.10 *Soit* $0 < a < d < b$, $r > 0$, $\lambda > 1$.
Sous $(\overline{HD}2)$*: il existe* $\epsilon_0 > 0$, $k > 0$ *tels que* $\forall \epsilon \in (0, \epsilon_0]$,

$$P\left(\left\{\frac{1}{\epsilon^2}\int_d^b \left|\hat{h}_t^+ - \hat{h}_t^-\right|^2 dt < r\right\} \cap C^\epsilon(a,b)\right) \leq \hat{\delta}_{[a,b]}\left(\frac{\lambda r}{b-d}\right) + e^{-k/\sqrt{\epsilon}}.$$

Preuve D'après la démonstration du lemme précédent dans la situation $(\overline{HD}2)$, il suffit de s'intéresser à

$$P\left(\left\{\frac{1}{\epsilon^2}\int_d^b \left|\epsilon\Lambda^{-1} D H_+ \hat{X}_t^+\right|^2 dt < \lambda r\right\} \cap C^\epsilon(a,b)\right)$$

En utilisant la définition (6.27) de la fonction $\hat{\delta}_{[a,b]}$, on conclut par les inégalités suivantes:

$$\begin{aligned}
& P\left(\left\{\frac{1}{\epsilon^2}\int_d^b \left|\epsilon\Lambda^{-1} D H_+ \hat{X}_t^+\right|^2 dt < \lambda r\right\} \cap C^\epsilon(a,b)\right) \\
\leq\ & P\left(\left\{\inf_{[a,b]} \left|\Lambda^{-1} D H_+ \hat{X}_t^+\right|^2 < \frac{\lambda r}{b-d}\right\} \cap C^\epsilon(a,b)\right) \\
\leq\ & \hat{\delta}_{[a,b]}\left(\frac{\lambda r}{b-d}\right).
\end{aligned}$$

□

Sous l'hypothèse $(\overline{HD}1)$, on a le résultat suivant:

Théorème 6.11 *Soit* $0 < a < d < b$.
Sous $(\overline{HD}1)$, *il existe* $\varepsilon_0 > 0$, $k > 0$, *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P(A_+^c(a,b) \mid \{\hat{L}^\varepsilon \geq 0\} \cap C^\varepsilon(a,b)) \leq e^{-k/\sqrt{\varepsilon}},$$
$$P(A_-^c(a,b) \mid \{\hat{L}^\varepsilon < 0\} \cap C^\varepsilon(a,b)) \leq e^{-k/\sqrt{\varepsilon}}.$$

Preuve On démontre la première inégalité:

$$A_+^c \cap \{\hat{L} \geq 0\} \cap C^\varepsilon(a,b) \subset [(A_+ \cup A_-)^c \cap C^\varepsilon] \cup [A_- \cap C^\varepsilon \cap \{\hat{L}^\varepsilon \geq 0\} \cap C^\varepsilon].$$

Le théorème 3.4 implique

$$P((A_+ \cup A_-)^c \cap C^\varepsilon) \leq e^{-k/\sqrt{\varepsilon}}.$$

D'après (6.11), sur l'événement A_-, la statistique $\hat{L}^\varepsilon$ s'écrit

$$\begin{aligned} \hat{L}^\varepsilon = & - \frac{1}{2}\langle N^- \rangle_b + N_b^- \\ & + \frac{1}{\varepsilon^2}\int_d^b (\hat{h}_t^+ - \hat{h}_t^-)(H_- \tilde{M}_t^- - H_- \hat{X}_t^-)\, dt. \end{aligned} \tag{6.28}$$

où le processus $\{N_t^-,\ t \geq 0\}$ est une martingale définie comme il suit

$$N_t^- = \frac{1}{\varepsilon}\int_d^t (\hat{h}_t^+ - \hat{h}_t^-) d\tilde{\omega}_t^-.$$

Les lemmes 6.6 et 6.7 dans le "cas négatif", nous indiquent que le dernier terme de l'expression (6.28) est négligeable devant le premier. En introduisant une martingale exponentielle et en utilisant le lemme 6.7, on obtient facilement le résultat. (Voir [4][th5.4]).

□

Soit $r > 0$ choisi, on définit les événements suivants:

$$\begin{aligned} \hat{C}^{\varepsilon,r}(a,b) &= C^\varepsilon \cap \left\{ \frac{1}{\varepsilon^2}\int_d^b (\hat{h}_t^+ - \hat{h}_t^-)^2 dt \geq r \right\} \\ \hat{C}_+^{\varepsilon,r}(a,b) &= \hat{C}^{\varepsilon,r}(a,b) \cap \{\hat{L}^\varepsilon \geq 0\} \\ \hat{C}_-^{\varepsilon,r}(a,b) &= \hat{C}^{\varepsilon,r}(a,b) \cap \{\hat{L}^\varepsilon < 0\} \end{aligned}$$

Proposition 6.12 *Soit* $0 < a < d < b$, $r > 0$.
Sous $(\overline{HD}2)$, *il existe* $\varepsilon_0 > 0$, $k > 0$, *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\begin{aligned} P(A_+^c(a,b) \cap \widehat{C}_+^{\varepsilon,r}(a,b)) &\leq e^{-k/\sqrt{\varepsilon}} + e^{-\frac{1}{32}r}, \\ P(A_-^c(a,b) \cap \widehat{C}_-^{\varepsilon,r}(a,b)) &\leq e^{-k/\sqrt{\varepsilon}} + e^{-\frac{1}{32}r}. \end{aligned}$$

Preuve On considère la seconde inégalité. D'après le théorème 3.4, nous considérons uniquement la probabilité

$$P(A_+(a,b) \cap \widehat{C}_-^{\varepsilon,r}(a,b)).$$

D'après (6.10), Sur l'événement A_+, la statistique $\widehat{L}^\varepsilon$ s'écrit:

$$\begin{aligned} \widehat{L}^\varepsilon = \quad & \frac{1}{4}\langle N^+ \rangle_b + N_b^+ \\ + \quad & \frac{1}{4}\langle N^+ \rangle_b + \frac{1}{\varepsilon^2}\int_d^b (\hat{h}_t^+ - \hat{h}_t^-)(H_+ \tilde{M}_t^+ - H_+ \hat{X}_t^+)\, dt, \end{aligned}$$

avec

$$N_t^+ = \frac{1}{\varepsilon}\int_d^t (\hat{h}_s^+ - \hat{h}_s^-) d\tilde{\omega}_s^+.$$

La somme des deux derniers termes est minorée par la variable aléatoire Z,

$$Z = \frac{1}{8}\langle N^+ \rangle_b - \frac{2|H_+|^2}{\varepsilon^2}\int_d^b (\tilde{M}_s^+ - \hat{X}_s^+)^2 ds.$$

Montrons que la variable Z est strictement positive avec une probabilité proche de 1. $\forall \theta > 0$,

$$\begin{aligned} P(\{Z < 0\} \cap A_+ \cap C^\varepsilon) \leq \quad & P\left(\left\{\frac{1}{8}\langle N^+ \rangle_b < \theta\right\} \cap C^\varepsilon\right) \\ & + P\left(\left\{\frac{2|H_+|^2}{\varepsilon^2}\int_d^b (\tilde{M}_s^+ - \hat{X}_s^+)^2 ds \geq \theta\right\} \cap A_+\right) \end{aligned}$$

On applique le lemme 6.7 au premier terme en choisissant $\theta \geq r_{inf}^2$. On utilise le lemme 6.6 pour majorer le second terme et on obtient

$$P(\{Z < 0\} \cap A_+ \cap C^\varepsilon) \leq e^{-k/\sqrt{\varepsilon}}.$$

pour ε suffisament petit. Considérons les deux premiers termes de la statistique $\widehat{L}^\varepsilon$. On doit estimer la quantité

$$P\Big(A_+ \cap \Big\{N_b + \tfrac{1}{4}\langle N\rangle_b < 0\Big\} \cap \{\langle N\rangle_b \geq r\}\Big).$$

Par le lemme 2.8, on sait que $\forall \mu \in \mathbb{R}$,

$$\tilde{\mathbf{E}}^+\left(\exp\left\{\mu N_b - \mu^2/2\langle N\rangle_b\right\}\right) = 1.$$

Sur $A_+ \cap \{\langle N\rangle_b \geq r\}$, si $\mu < 0$, on a la minoration suivante:

$$\mu N_b - \frac{\mu^2}{2}\langle N\rangle_b \geq -\frac{\mu}{2}(\frac{1}{2} + \mu)r.$$

En prennant $\mu = -\frac{1}{4}$, on obtient

$$\begin{aligned} & P\left(A_+ \cap \left\{\widehat{L}^\varepsilon < 0\right\} \cap \{\langle N\rangle_b \geq r\}\right) \\ \leq\ & \tilde{P}^+\left(\exp\left\{-\frac{1}{4}N_b - \frac{1}{32}\langle N\rangle_b\right\} \geq e^{\frac{1}{32}r}\right) \leq e^{-\frac{r}{32}}. \end{aligned}$$

□

Montrons le théorème suivant:

Théorème 6.13 *Soit* $0 < a < d < b$, $\lambda > 1$, $r_0 > 0$.

Sous $(\overline{HD}2)$, *il existe* k, $\varepsilon_0 > 0$, *une fonction croissante* $\hat{\rho}^\varepsilon_{[a,b]} : \mathbb{R}^+ \to \mathbb{R}$ *tels que* $\forall r \in (0, r_0]$, $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P(A^c_+(a,b) \,|\, \widehat{C}^{\varepsilon,r}_+(a,b)) \leq e^{-k/\sqrt{\varepsilon}} + \hat{\rho}^\varepsilon_{[a,b]}\left(\frac{\lambda r}{b-d}\right) e^{-\frac{1}{32}r},$$

$$P(A^c_-(a,b) \,|\, \widehat{C}^{\varepsilon,r}_-(a,b)) \leq e^{-k/\sqrt{\varepsilon}} + \hat{\rho}^\varepsilon_{[a,b]}\left(\frac{\lambda r}{b-d}\right) e^{-\frac{1}{32}r},$$

Preuve D'après la proposition 6.12, il nous reste à minorer les probabilités $P(\widehat{C}^{\varepsilon,r}_+(a,b))$ et $P(\widehat{C}^{\varepsilon,r}_-(a,b))$. La démonstration est similaire à celle du théorème 5.13. La fonction $\hat{\rho}^\varepsilon_{[a,b]} : \mathbb{R}^+ \to \mathbb{R}$, est définie par

$$\hat{\rho}^\varepsilon_{[a,b]}(x) = \frac{1}{\beta - \hat{\delta}_{[a,b]}(x) - e^{-k/\sqrt{\varepsilon}}}, \tag{6.29}$$

où la fonction $\hat{\delta}_{[a,b]} : \mathbb{R}^+ \to [0,1[$ est donnée par (6.27).

□

Remarque 6.14

$$\forall x \geq 0, \quad \hat{\rho}^{\varepsilon}_{[a,b]}(x) \geq \hat{\rho}^{\varepsilon}_{[a,b]}\left(\frac{r^2_{inf}}{b-d}\right) = \frac{1}{\beta - e^{-k/\sqrt{\varepsilon}}}.$$

Comme le test présenté dans le paragraphe 5, sous $(\overline{HD2})$, ce test ne pourra décider "correctement" du signe de $\langle X_t, u\rangle$ que sur des intervalles de monotonie suffisament longs. (Voir la remarque 5.15).

Procédure du test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy:

De même que pour les tests précédents, on introduit τ_1, $a < d < \tau_1 \leq b$, le temps à partir duquel le test peut prendre une décision pour une probabilité d'erreur donnée. On définit respectivement sous $(\overline{HD1})$ et sous $(\overline{HD2})$ deux nouveaux événements tests :

- Sous $(\overline{HD1})$,

$$\hat{Q}^{\varepsilon}_{+}(a,b) = C^{\varepsilon}(a,b) \cap \{\hat{L}^{\varepsilon}(a,\tau_1) \geq 0\},$$

$$\hat{Q}^{\varepsilon}_{-}(a,b) = C^{\varepsilon}(a,b) \cap \{\hat{L}^{\varepsilon}(a,\tau_1) < 0\},$$

- Sous $(\overline{HD2})$,

$$\hat{Q}^{\varepsilon,r}_{+}(a,b) = C^{\varepsilon}(a,b) \cap \hat{C}^{\varepsilon,r}_{-}(a,\tau_1),$$
$$\hat{Q}^{\varepsilon,r}_{-}(a,b) = C^{\varepsilon}(a,b) \cap \hat{C}^{\varepsilon,r}_{-}(a,\tau_1).$$

A l'aide des théorèmes 3.4, 6.11 et 6.13, on déduit par un raisonnement identique à celui du corollaire 5.16, le résultat suivant:

Corollaire 6.15 *Soit* $0 < a < d < \tau_1 \leq b$.

Sous $(\overline{HD1})$,
il existe $\varepsilon_0 > 0$, $k > 0$, *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P(A^{\varepsilon}_{+}(a,b) \mid \hat{Q}^{\varepsilon}_{+}(a,b)) \leq e^{-k/\sqrt{\varepsilon}},$$
$$P(A^{\varepsilon}_{-}(a,b) \mid \hat{Q}^{\varepsilon}_{-}(a,b)) \leq e^{-k/\sqrt{\varepsilon}}.$$

Sous $(\overline{HD}2)$,
soit $\lambda > 1$, $r_0 > 0$, *il existe* k, $\varepsilon_0 > 0$, *une fonction croissante* $\hat{\rho}^{\varepsilon}_{[a,b]}$: $\mathbb{R}^+ \to \mathbb{R}$ *tels que* $\forall r \in (0, r_0]$, $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P(A^c_+(a,b) \,|\, \widehat{Q}^{\varepsilon,r}_+(a,b)) \leq e^{-k/\sqrt{\varepsilon}} + \hat{\rho}^{\varepsilon}_{[a,b]}\left(\frac{\lambda r}{\tau_1 - d}\right) e^{-\frac{1}{32}r},$$

$$P(A^c_-(a,b) \,|\, \widehat{Q}^{\varepsilon,r}_-(a,b)) \leq e^{-k/\sqrt{\varepsilon}} + \hat{\rho}^{\varepsilon}_{[a,b]}\left(\frac{\lambda r}{\tau_1 - d}\right) e^{-\frac{1}{32}r},$$

□

Ce corollaire met en évidence que le comportement de ce test diffère sous l'hypothèse $(\overline{HD}1)$ et $(\overline{HD}2)$. Dans la situation du type (2), il ne peut prendre des décisions correctes que sur des intervalles de temps suffisament grands. Dans les cas où les hypothèses $(\overline{HD}1)$ et $(HD2)$ sont vérifiées simultanément(voir exemple 6.3), il est clair qu'il est préférable d'utiliser le test sur les sorties des filtres de Kalman-bucy plutôt que celui sur les accroissements de l'observation présenté dans le paragraphe 5.

Il est important de souligner qu'en dimension $n > 1$, dans certains cas, le test sur les sorties de Kalman-Bucy permet de décider du signe de $\langle X_t, u \rangle$ dans une situation du type (1) alors que le test de la variation quadratique ne peut être appliqué.

Remarque 6.16 Ce test a été mis en œuvre sous forme séquentielle dans le chapitre II. Comme dans le test du rapport de vraisemblance sur les accroissements de l'observation (voir remarque 5.17), on détermine, pour une probabilité d'erreur donnée, des bornes $l_1 > 0$ et $l_2 > 0$ et non la borne r du terme de précision. elles permettront de prendre une décision dès que $L^{\varepsilon}(a, \tau_1) \leq -l_1$ ou $L^{\varepsilon}(a, \tau_1) \geq l_2$. Ces bornes sont différentes sous les deux hypothèses $(\overline{HD}1)$ et $(\overline{HD}2)$.

7 Erreurs d'approximation

Supposons, que l'on ait conclu à l'aide d'une des procédures présentées précédemment que $X_t > 0$ pour $t \in [a, b]$. Nous prendrons alors $\hat{X}_b^+$ comme approximation de $\mathbf{E}(X_b \mid \mathcal{Y}_b)$. Estimons l'erreur d'approximation de la moyenne conditionnelle. Soit $0 < a < \tau_1 \leq b$, on considère les événements suivants:

sous $(HD1)$:

$$\begin{aligned} Q_+^\epsilon(a,b) &= C^\epsilon(a,b) \cap C_+^\epsilon(a,\tau_1), \\ Q_-^\epsilon(a,b) &= C^\epsilon(a,b) \cap C_-^\epsilon(a,\tau_1). \end{aligned}$$

Sous $(HD2)$:

$$\begin{aligned} Q_+^{\epsilon,r}(a,b) &= C^\epsilon(a,b) \cap C_+^{\epsilon,r}(a,\tau_1), \\ Q_-^{\epsilon,r}(a,b) &= C^\epsilon(a,b) \cap C_-^{\epsilon,r}(a,\tau_1). \end{aligned}$$

Sous les deux hypothèses $(\overline{HD1})$ et $(\overline{HD2})$, la discussion avec les événements du test du rapport de vraisemblance sur les sorties des filtres de Kalman est similaire.

On note $\hat{X}_t = \mathbf{E}(X_t \mid \mathcal{Y}_t)$. Nous voulons estimer:

Sous (HD1) : $|\hat{X}_b - \hat{X}_b^+|$ sur Q_+^ϵ.

Sous (HD2) : $|\hat{X}_b - \hat{X}_b^+|$ sur $Q_+^{\epsilon,r}$.

L'erreur d'approximation dans le "cas négatif" s'obtient de la même façon.

On introduit la décomposition suivante:

$$\hat{X}_b - \hat{X}_b^+ = (\hat{X}_b - \hat{M}_b^+) + (\hat{M}_b^+ - \tilde{M}_b^+) + (\tilde{M}_b^+ - \hat{X}_b^+), \qquad (7.1)$$

où le processus intermédiaire $\{M_t^+, t \geq 0\}$ est défini dans le paragraphe précédent par (6.5). Nous utiliserons les deux résultats suivants:

Lemme 7.1 *Il existe $\varepsilon_0 > 0$, $k > 0$, tels que $\forall \varepsilon \in (0, \varepsilon_0]$,*

$$\hat{P}^+(A_+^c \cap C^\epsilon) \leq e^{-k/\epsilon}.$$

Preuve La démontration est similaire à celle du théorème 3.4.

1ière étape: On montre par un raisonnement identique à celui du lemme 3.1 que pour tout $\theta > 0$, il existe $k > 0$ tel que

$$\tilde{P}^+\left(\left\{\sup_{[a,b]}|M_s^+ - \hat{X}_s^+| > \theta\right\}\right) \leq e^{-k/\epsilon}. \tag{7.2}$$

2ième étape: Soit la constante $c > 0$ choisie pour définir C^ϵ.

$$\begin{aligned}\tilde{P}^+(A_+^c \cap C^\epsilon) \leq \quad & \tilde{P}^+\left(\left\{\sup_{[a,b]}|M_s^+ - \hat{X}_s^+| > c\right\}\right) \\ + \quad & \tilde{P}^+\left(\left\{\sup_{[a,b]}|M_s^+ - \hat{X}_s^+| \leq c\right\} \cap A_+^c \cap C^\epsilon\right)\end{aligned}$$

D'après (7.2), il suffit de s'intéresser au dernier terme. On introduit τ, un $\mathcal{F}_t$–temps d'arrêt,

$$\tau = \inf\{t \geq a\,,\, X_t^1 \leq 0\} \wedge T,$$

avec $X_t^1 = \langle X_t\,,\, u\rangle$. Sur l'événement A_+^c, $\tau \leq b$ et $\langle M_\tau^+\,,\, u\rangle = 0$, d'où

$$\left\{\sup_{[a,b]}|M_s^+ - \hat{X}_s^+| \leq c\right\} \cap A_+^c \cap C^\epsilon \subset \left\{\left|\hat{X}_\tau^+\right| \leq c\right\} \cap C^\epsilon = \emptyset$$

□

Lemme 7.2

Sous (HD1) : il existe ε_0, $k > 0$ tels que $\forall \epsilon \in (0, \varepsilon_0]$,

$$\mathbf{E}\left(|\widehat{M_b^+ Z_b} - \hat{M}_b^+|\,;\, Q_+^\epsilon\right) \leq e^{-k/\sqrt{\epsilon}}$$

Sous (HD2) : soit $r > 0$, il existe ε_0, k_1,k_2, $k_3 > 0$ tels que $\forall \epsilon \in (0, \varepsilon_0]$,

$$\mathbf{E}\left(|\widehat{M_b^+ Z_b} - \hat{M}_b^+|\,;\, Q_+^{\epsilon,r}\right) \leq e^{-k_1/\sqrt{\epsilon}} + k_2\, e^{-k_3 r}.$$

Preuve Si $\omega \in A_+$, $Z_b(\omega) = 1$.
Soit $D \in \mathcal{Y}_b$.

$$\begin{aligned}\mathbf{E}\left(|\widehat{M_b^+ Z_b} - \hat{M}_b^+|\,;\, D\right) &= \int_D |\widehat{M_b^+ Z_b} - \hat{M}_b^+| dP \\ &\leq \int_{D\cap A_+^c} |M_b^+ Z_b - M_b^+| dP \\ &\leq \int_{D\cap A_+^c} |M_b^+| d\tilde{P}^+ + \int_{D\cap A_+^c} |M_b^+| dP\end{aligned}$$

Par l'inégalité de Cauchy-Schwartz, on obtient

$$\begin{aligned}\mathbf{E}\left(|\widehat{M_b^+ Z_b} - \hat{M}_b^+|\,;\, D\right) \leq \quad & (\tilde{E}^+ \left|M_b^+\right|^2)^{\frac{1}{2}} \tilde{P}^+(D \cap A_+^c)^{\frac{1}{2}} \\ + \quad & (E\left|M_b^+\right|^2)^{\frac{1}{2}} P(D \cap A_+^c)^{\frac{1}{2}}. \qquad (7.3)\end{aligned}$$

On rappelle que le processus $\{M_t^+\,;\, t \geq 0\}$ vérifie les systèmes suivants:

$$\begin{cases} dM_t^+ &= B_+ M_t^+ dt + B_+(-\psi_t^a - \varphi_t^a + \phi_t^a)dt + F_+ dV_t + G_+ dW_t\,,\ t \geq a; \\ M_t^+ &= X_t^+\,,\ t \leq a \end{cases}$$

où $\{V_t\,,\ t \leq 0\}$ et $\{W_t\,,\ t \leq 0\}$ sont deux P–processus de Wiener standards indépendants.

$$\begin{cases} dM_t^+ &= B_+ M_t^+ dt + F_+ d\tilde{V}_t^+ + G_+ d\tilde{W}_t^+\,,\ t \geq a; \\ M_t^+ &= X_t^+\,,\ t \leq a. \end{cases}$$

où $\{\tilde{V}_t^+\,,\ t \leq 0\}$ et $\{\tilde{W}_t^+\,,\ t \leq 0\}$ sont deux $\tilde{P}^+$–processus de Wiener standards indépendants.

Par les propriétés du temps local (3.4), les lemmes 2.5 et 2.7, et l'inégalité de Cauchy-Schwartz, on obtient

$$\mathbf{E}(|M_b^+|^2) < +\infty\,,\ \tilde{E}^+(|M_b^+|^2) < +\infty.$$

Sous $(HD1)$, nous prennons $D = Q_+^c$, l'inégalité (7.3), le lemme 7.1 et le théorème 4.5 impliquent

$$\mathbf{E}\left(|\widehat{M_b^+ Z_b} - \hat{M}_b|\,;\, D\right) \leq e^{-k/\sqrt{\epsilon}},$$

pour un certain $k > 0$ et ε suffisament petit.

Sous $(HD2)$, nous prennons $D = Q_+^{\varepsilon,r}$. Par l'inégalité (7.3), le lemme 7.1 et la proposition 5.10, nous obtenons

$$\mathbf{E}\left(|\widehat{M_b^+ Z_b} - \hat{M}_b|; D^r\right) \leq e^{-k_1/\sqrt{\varepsilon}} + k_2\, e^{-k_3 r},$$

pour certains $k_1, k_2, k_3 > 0$ et ε suffisament petit.

□

Par une démonstration analogue à la précédente, on démontre le lemme suivant:

Lemme 7.3

Sous $(HD1)$: il existe $\varepsilon_0 > 0$, $k > 0$, tels que

$$\mathbf{E}\left(\{|\hat{Z}_b - 1|; Q_+^{\varepsilon}\right) \leq e^{-k/\sqrt{\varepsilon}}$$

Sous $(HD2)$: soit $r > 0$, il existe $\varepsilon_0 > 0$, $k_1 > 0$, $k_2 > 0$, $k_3 > 0$, tels que

$$\mathbf{E}\left(\{|\hat{Z}_b - 1|; Q_+^{\varepsilon,r}\right) \leq e^{-k_1/\sqrt{\varepsilon}} + k_2\, e^{-k_3 r}.$$

□

Théorème 7.4

Sous $(HD1)$: $\forall \theta > 0$, $q > 0$, il existe $\varepsilon_q > 0$, $k_q > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_q]$,

$$P(\{|\hat{X}_b - \hat{X}_b^+| > \theta \varepsilon^q\} \cap Q_+^{\varepsilon}) \leq e^{-k_q/\sqrt{\varepsilon}}.$$

Sous $(HD2)$: $\forall \theta > 0$, il existe $\varepsilon_0 > 0$, $k_1, k_2, k_3 > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P(\{|\hat{X}_b - \hat{X}_b^+| > \theta\} \cap Q_+^{\varepsilon,r}) \leq e^{-k_1/\sqrt{\varepsilon}} + \frac{k_2 r}{\theta} e^{-k_3 r}.$$

Preuve Nous considérons uniquement le cas où l'hypothèse $(HD2)$ est vérifiée. Dans l'autre cas, la démonstration est une version simplifiée de celle présentée. On introduit la décomposition suivante:

$$\hat{X}_b - \hat{X}_b^+ = (\hat{X}_b - \hat{M}_b^+) + (\hat{M}_b^+ - \tilde{M}_b^+) + (\tilde{M}_b^+ - \hat{X}_b^+).$$

1) $\hat{X}_b - \hat{M}_b^+$: En utilisant l'inégalité de Jensen et le fait que $X_b = M_b^+$ sur l'événement A_+, on obtient:

$$\begin{aligned}\mathbf{E}(|\hat{X}_b - \hat{M}_b^+|\,;\, Q_+^{\epsilon,r}) &\leq \mathbf{E}(|X_b - M_b^+|\,;\, Q_+^{\epsilon,r} \cap A_+^c) \\ &\leq \mathbf{E}(|X_b - M_b^+|^2)^{\frac{1}{2}} P(Q_+^{\epsilon,r} \cap A_+^c)^{\frac{1}{2}}\end{aligned}$$

Le moment $\mathbf{E}|X_b - M_b^+|^2$ étant borné par une constante indépendante de ϵ, donc l'inégalité (5.15) implique l'existence de k_1, k_2 et $k_3 > 0$ tels que pour ϵ suffisament petit,

$$P(\{|\hat{X}_b - \hat{M}_b^+| > \theta\} \cap C_+^{\epsilon,r}) \leq e^{-k_1/\sqrt{\epsilon}} + \frac{k_2}{\theta} e^{-k_3 r} \tag{7.4}$$

2) $\hat{M}_b^+ - \tilde{M}_b^+$: cette différence se réécrit

$$\hat{M}_b^+ - \tilde{M}_b^+ = \hat{M}_b^+ - \tilde{M}_b^+ \hat{Z}_b + \tilde{M}_b^+ (\hat{Z}_b - 1).$$

En utilisant l'égalité $\tilde{M}_b^+ \hat{Z}_b = \widehat{M_b^+ Z_b}$, on obtient:

$$\begin{aligned}& P(\{|\hat{M}_b^+ - \tilde{M}_b^+| > \theta\} \cap Q_+^{\epsilon,r}) \\ \leq\ & P\left(\left\{|\widehat{M_b^+ Z_b} - \hat{M}_b^+| > \frac{\theta}{2}\right\} \cap Q_+^{\epsilon,r}\right) \\ +\ & P\left(\left\{|\tilde{M}_b^+|\,|\hat{Z}_b - 1| > \frac{\theta}{2}\right\} \cap Q_+^{\epsilon,r}\right).\end{aligned}$$

Le premier terme se traite par l'inégalité de Markov et le lemme 7.2. On s'intéresse au second terme. Soit $\beta > 0$,

$$\begin{aligned}& P\left(\left\{|\tilde{M}_b^+|\,|\hat{Z}_b - 1| > \frac{\theta}{2}\right\} \cap Q_+^{\epsilon,r}\right) \\ \leq\ & P(Q_+^{\epsilon,r} \cap A_+^c) + \tilde{P}\left(\left\{|\tilde{M}_b^+| > \frac{1}{\beta}\right\} \cap A_+ \cap Q_+^{\epsilon,r}\right) \\ +\ & P\left(\left\{|\hat{Z}_b - 1| > \frac{\beta\theta}{2}\right\} \cap Q_+^{\epsilon,r}\right) \\ \leq\ & e^{-k_1/\sqrt{\epsilon}} + k_2\, e^{-k_3/\beta} + \frac{k_4}{\theta\beta} e^{-k_5 r},\end{aligned}$$

pour certains $k_1, k_2, k_3, k_4, k_5 > 0$ et ε assez petit, en utilisant respectivement pour chaque terme, l'inégalité (5.15), le lemme 2.5 et le lemme 7.3. Si nous prennons $\beta = \frac{1}{r}$, on obtient

$$\tilde{P}(\{|\hat{M}_b^+ - \tilde{M}_b^+| > \theta\} \cap Q_+^{\varepsilon,r}) \leq e^{-k_1/\sqrt{\varepsilon}} + \frac{k_2 r}{\theta} e^{-k_3 r}. \tag{7.5}$$

pour certains $k_1, k_2, k_3 > 0$ et ε petit.

3) $\tilde{M}_b^+ - \hat{X}_b^+$: D'après l'inégalité (6.16), il existe ε_0, $k > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\tilde{E}^+ |\tilde{M}_b^+ - \hat{X}_b^+| < e^{-k/\varepsilon} \tag{7.6}$$

Ainsi, d'après l'inégalité (5.15) et (7.6), on obtient

$$\begin{aligned} P(\{|\tilde{M}_b^+ - \hat{X}_b^+| > \theta\} \cap Q_+^{\varepsilon,r}) \leq & \quad P(Q_+^{\varepsilon,r} \cap A_+^c) \\ & + \tilde{P}^+(\{|\tilde{M}_b^+ - \hat{X}_b^+| > \theta\} \cap A_+) \\ \leq & \quad e^{-k_1/\sqrt{\varepsilon}} + \frac{k_2}{\theta} e^{-k_3 r}. \end{aligned} \tag{7.7}$$

Des inégalités (7.4), (7.5) et (7.7), on déduit le théorème sous l'hypothèse $(HD2)$.

□

Sous l'hypothèse $(HD1)$, d'après le théorème précédent, le filtre de Kalman à l'instant b, $\hat{X}_b^+$, est asymptotiquement optimal sur l'événement-test $Q_+^\varepsilon(a, b)$ lorsque $\varepsilon \to 0$.

Corollaire 7.5 *Sous* $(HD1)$, $\forall q > 0$, *il existe* ε_q, $k_q = k_q(\varepsilon_q) > 0$ *tels que*

$$\frac{1}{\varepsilon^q} \mathbf{E}\left(|\hat{X}_b - \hat{X}_b^+|^q ; Q_+^\varepsilon\right) \leq k_q$$

□

Du théorème 7.4 et du lemme 5.11, on déduit le corollaire suivant:

Corollaire 7.6 *Soit* $0 < a < \tau_1 \leq b$,

Sous $(HD1)$: $\forall\theta,\ q>0$, *il existe* $\varepsilon_q,\ k_q>0$ *tels que* $\forall\varepsilon\in(0,\varepsilon_q]$,

$$P(\{|\hat{X}_b-\hat{X}_b^+|>\theta\varepsilon^q\}\,|\,Q_+^\varepsilon)\le e^{-k_q/\sqrt{\varepsilon}}.$$

Sous $(HD2)$: $\forall\theta>0,\ \lambda>1$, *il existe* $r_0>0,\ \varepsilon_0>0,\ k_1,k_2,k_3>0$ *et une fonction croissante* $\rho^\varepsilon_{[a,b]}:\mathbb{R}^+\to\mathbb{R}^+$, *tels que* $\forall r\in(0,r_0]$, $\forall\varepsilon\in(0,\varepsilon_0]$,

$$P(\{|\hat{X}_b-\hat{X}_b^+|>\theta\}\,|\,Q_+^{\varepsilon,r})\le e^{-k_1/\sqrt{\varepsilon}}+\frac{k_2 r}{\theta}\,\rho^\varepsilon_{[a,b]}\left(\frac{\lambda r}{\tau_1-a}\right)e^{-k_3 r}.$$

On a un résultat similaire pour l'erreur d'approximation $|\hat{X}_b-\hat{X}_b^-|$ *sur* Q_-^ε *sous* $(HD1)$, *respectivement, sur* $Q_-^{\varepsilon,r}$ *sous* $(HD2)$.

□

8 Extension des résultats au cas $l \geq 2$

Nous étudions brièvement comment adapter les différents tests présentés dans les paragraphes 3, 4, 5 et 6 dans le cas où la fonction d'observation h est affine par "morceaux" dans le cas $l \geq 2$. Pour éviter que le processus $\{X_t\,;\, t \geq 0\}$ puisse atteindre les sous espaces de $\mathbb{R}^n$ de dimension au plus $n-2$, (cf. [1]), nous considérons uniquement le cas où le coefficient de diffusion est constant.

Les hypothèses $(H1)$ à $(H5)$ sont supposées vérifiées avec $\Sigma_1 = \Sigma_2 = \ldots = \Sigma_l = I_n$. On rappelle que nous supposons l'existence d'une partition polyédrale finie de $\mathbb{R}^n$,

$$\mathbb{R}^n = \bigcup_{i=1}^{l} \Theta_i$$

telle que sur chaque polyèdre Θ_i, les restrictions des applications continues b et h sont affines et celles de l'application mesurable σ sont constantes. Pour $i = 1, \ldots, l$, on définit les applications suivantes:

$$\begin{aligned} \bar{b}_i : \mathbb{R}^n &\to \mathbb{R}^n \\ x &\to B_i x + b_i\,, \end{aligned}$$

$$\begin{aligned} \bar{h}_i : \mathbb{R}^n &\to \mathbb{R}^n \\ x &\to H_i x + h_i\,. \end{aligned}$$

Ainsi, $\forall x \in \Theta_i$, on a

$$b(x) = \bar{b}_i(x)\,,\ h(x) = \bar{h}_i(x) \text{ et } \sigma(x) = I_n.$$

On note que l'hypothèse de continuité des applications h et b implique que cette partition est "propre" i.e. la frontière entre deux ployèdres est incluse soit dans un sous espace de $\mathbb{R}^n$ de dimension au plus $n-2$, soit dans un hyperplan Δ_j,

$$\Delta_j = \{x \in \mathbb{R}^n\,;\, \langle x\,,\, u_j \rangle = c_j\},$$

avec u_j un vecteur normé de $\mathbb{R}^n$ et $c_j \in \mathbb{R}^n$.

Nous désignerons respectivement l'intérieur, l'adhérence et la frontière d'un polyèdre Θ par $\mathring{\Theta}$, $\overline{\Theta}$ et $\partial\Theta$. La réunion $\mathcal{U}$ de toutes les frontières des polyèdres $\{\Theta_i\}_{j=1,l}$ peut s'écrire sous la forme suivante:

$$\mathcal{U} = \bigcup_{j=1}^{J} D_j \cup \mathcal{N}\,,$$

où $D_j = \{x \in \mathbb{R}^n,\ \langle x, u_j\rangle = c_j$ et $\alpha_j < \langle x, v_j\rangle < \beta_j\}$ avec u_j, v_j des vecteurs normés de $\mathbb{R}^n$, c_j, α_j et β_j des réels, est inclus dans un hyperplan Δ_j et $\mathcal{N}$ est la réunion des sous espaces de dimension au plus $n-2$.

Soit $0 < a < b$, on définit les événements suivants:

$$A_i(a,b) = \{X_t \in \mathring{\Theta}_i,\ a \le t \le b\} \quad i = 1,\dots,l.$$

Pour $c > 0$ choisi,

$$C^c(a,b) = \{\ \mathrm{dist}(\bar{h}_i(\hat{X}_t^i),\ h(\mathcal{U})) > c,\ a \le t \le b\}.$$

Test de détection des passages de $\{X_t\}$ par $\mathcal{U}$:

On s'intéresse à la différence $h(X_t) - \bar{h}_i(\hat{X}_t^i)$, $i = 1,\dots,l$. Vu que nous avons restreint notre étude au cas où le processus $\{X_t\,;\ t \ge 0\}$ a une probabilité nulle d'atteindre $\mathcal{N}$, nous pouvons utiliser la formule d'Itô-Tanaka généralisée pour obtenir la différentielle de $h(X_t)$ (cf. [11]).

Soit $x \in D_j$, on a, pour ε petit,

$$h(x - \varepsilon u_j) = H_j^-(x - \varepsilon u_j) + h_j^-,$$
$$h(x + \varepsilon u_j) = H_j^+(x + \varepsilon u_j) + h_j^+.$$

On pose $(\Delta H)_j = (H_j^- - H_j^+)$.

$$dh(X_t) = h'(X_t)dX_s + \frac{1}{2}\sum_{j=1}^{J}(\Delta H)_j u_j \mathbf{1}_{\{X_t \in D_j\}} dL_t^{c_j}(\langle X\ ,\ u_j\rangle),$$

où $\{L_t^{c_j}\,;\ t \ge 0\}$ est le processus temps local en c_j de la semi-martingale réelle $\langle X\ ,\ u_j\rangle$ et h' est la dérivée de h définie sur $\mathbb{R}^n - \mathcal{U}$.

D'autre part,

$$d\bar{h}_i(\hat{X}_t^i) = H_i(B_i\hat{X}_t^i + b_i)dt + \frac{1}{\varepsilon}H_i K_\infty^i[h(X_t) - \bar{h}_i(\hat{X}_t^i)]dt + H_i K_\infty^i dW_t.$$

Donc $\{Z_t = h(X_t) - \bar{h}_i(\hat{X}_t^i)\,;\ t \ge 0\}$, vérifie l'équation suivante:

$$dZ_t = \frac{1}{\varepsilon}\Lambda_i Z_t dt + \varphi_t^i dt + \psi_t^{1,i} dV_t + \psi_t^{2,i} dW_t + \sum_{j=1}^{J}(\Delta H)_j u_j \mathbf{1}_{\{X_t \in D_j\}} dL_t^{c_j}(\langle X\ ,\ u_j\rangle), \tag{8.1}$$

avec

- $\Lambda_i = -H_i K_\infty^i$, matrice stable d'après la remarque (2.2).
- $\varphi_t^i = h'(X_t)b(X_t) - H_i(B_i\hat{X}_t^i + b_i)$.
- $\psi_t^1 = h'(X_t)f(X_t)$.
- $\psi_t^2 = h'(X_t)g(X_t) + H_i K_\infty^i$.

A partir de (8.1), par une démonstration similaire à celle de la proposition 3.1, on établit le résultat suivant:

Proposition 8.1 $\forall i = 1, \ldots, l$, $\forall \theta > 0$, $\forall \alpha \in [0, \frac{1}{2}[$, $\forall \gamma \in]2\alpha, 1[$, *il existe* $k > 0$, $\varepsilon_0 > 0$ *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\sup_{[a,b]} \left|h(X_t) - \bar{h}_i(\hat{X}_t^i)\right| > \theta\varepsilon^\alpha\right\}\right) \leq \exp\left\{-\frac{k}{\varepsilon^{\gamma-2\alpha}}\right\}.$$

□

Ceci nous permet de construire un test de détection des passages de X_t par $\mathcal{U}$ à l'aide de $\hat{X}_t^i$. Choisissons par exemple $i = l$. Le résultat analogue à celui du théorème 3.4 s'écrit:

$$P(\{(\bigcup_{i=1}^{l} A_i)^c \cap C^\varepsilon\} \leq e^{-k/\sqrt{\varepsilon}},$$

pour un certain $k > 0$ et ε petit.

Remarque 8.2 Appliqué dans un cadre général, ce test peut être très sélectif lorsque l est grand, au sens où il détecte plus de passages de X_t par $\mathcal{U}$ qu'il n'y en a. En effet, dans le cas $l > 2$, du fait de la non injectivité de h, l'image de la frontière d'un polyèdre peut être incluse dans l'image de parties de $\mathbb{R}^n$ d'intersection vide avec $\mathcal{U}$. Ainsi, pour l grand, ce test est intéressant dans sa version locale, lorsque les valeurs prises par la fonction h permettent de découper le domaine d'application du test.

Pour $p \geq 2$, supposons qu'il existe $k_1, k_2, \ldots, k_p$ avec $1 \leq k_1 < k_2 < \cdots < k_p \leq l$ tel que

$$\bigcap_{i=1}^{p} h(\Theta_{k_i}) \neq \emptyset$$

et un compact $\mathcal{K}$ inclus dans $\bigcap_{i=1}^{p} h(\Theta_{k_i})$ tel que

$$0 < c < \operatorname{dist}(h(\mathcal{U}), \mathcal{K}).$$

On définit l'événement C_{ij}^{ε},

$$C_{\mathcal{K}}^{\varepsilon} = \{\hat{h}_l(\hat{X}_t^l) \in \mathcal{K}\,,\ \forall t \in [a,b]\}.$$

- Sur $C_{\mathcal{K}}^{\varepsilon}$, on effectue un premier choix entre "A_{k_1}" et "A_{k_2}".
- Si, par exemple, le test décide "A_{k_1}", sur $C_{\mathcal{K}'}^{\varepsilon}$ on teste "A_{k_1}" contre "A_{k_3}".
- ....
- On réitère la procédure jusqu'à ce que toutes les alternatives soient considérées.

Cependant, si en pratique un des tests ne peut décider faute de temps sur l'intervalle $[a,b]$, aucune décision ne pourra être prise sur cet intervalle. Ainsi, il est suffisant de décrire les tests de choix entre deux alternatives "A_i" et "A_j".

Soit $\mathcal{K}_{ij}$, un compact inclus dans $h(\Theta_i) \cap h(\Theta_j)$ tel que

$$0 < c < \operatorname{dist}(h(\mathcal{U}), K_{ij}).$$

Sur l'événement $C_{ij}^{\varepsilon} = \{\hat{h}_l(\hat{X}_t^l) \in O_{ij}\,;\ \forall t \in [a,b]\}$, nous devons appliquer un test pour choisir entre "A_i" et "A_j".

Sous $(HD1)$, le test de la variation quadratique:

L'hypothèse $(HD1)$ s'écrit:

$$H_i\Sigma_i\Sigma_i^*H_i^* \neq H_j\Sigma_j\Sigma_j^*H_j^*$$

On définit les événements $Q_i^{\varepsilon} = Q_+^{\varepsilon}$, $Q_j^{\varepsilon} = Q_-^{\varepsilon}$ comme en (4.1) et (4.2). Alors pour ε petit,

$$P(A_i \mid Q_j^{\varepsilon}) \leq e^{k/\sqrt{\varepsilon}},$$
$$P(A_j \mid Q_i^{\varepsilon}) \leq e^{k/\sqrt{\varepsilon}}.$$

Sous ($HD2$), test du rapport de vraisemblance sur les accroissements de l'observation:

Sous l'hypothèse ($HD2$),

$$\begin{cases} \Gamma\Gamma^* = \Gamma_i\Gamma_i^* = \Gamma_j\Gamma_j^*, \\ \ker\left(H_i \circ \bar{b}_i \circ \bar{h}_i^{-1} - H_j \circ \bar{b}_j \circ \bar{h}_j^{-1}\right) \subset h(\mathcal{U}), \\ \text{la matrice } (\Gamma\Gamma^*)^{-1}(H_iB_iH_i^{-1} - H_jB_jH_j^{-1}) \text{ est symétrique,} \end{cases}$$

à partir de l'étude du cas $\epsilon = 0$, nous définissons la fonction suivante:

$$\begin{array}{rccl} \kappa_{ij}: & \mathbb{R}^n & \to & \mathbb{R} \\ & x & \to & \frac{1}{2}x^*(\Gamma\Gamma^*)^{-1}\left(H_iB_iH_i^{-1} - H_jB_jH_j^{-1}\right)x \\ & & & +(\Gamma\Gamma^*)^{-1}\left(H_i(b_i - B_iH_i^{-1}h_i) - H_j(b_j - B_jH_j^{-1}h_j)\right)x \end{array}$$

et la statistique L_{ij}^ϵ,

$$\begin{aligned} L_{ij}^\epsilon &= \kappa_{ij}(\bar{y}_{m-1}) - \kappa_{ij}(\bar{y}_0) \\ &- \frac{1}{2}(b-a)\ \text{Trace}\ (H_iB_iH_i^{-1} - H_jB_jH_j^{-1}) \\ &- \frac{\epsilon}{2}\sum_{k=0}^{m-1}\left(|\Gamma^{-1}H_i \circ \bar{b}_i \circ \bar{h}_i^{-1}(\bar{y}_k)|^2 - |\Gamma^{-1}H_j \circ \bar{b}_j \circ \bar{h}_j^{-1}(\bar{y}_k)|^2\right), \end{aligned}$$

en utilisant les notations introduites dans le paragraphe 4. Le terme de précision R_{ij}^ϵ s'écrit

$$R_{ij}^\epsilon = \epsilon\sum_{k=0}^{m-1}|\Gamma^{-1}(H_i \circ \bar{b}_i \circ \bar{h}_i^{-1} - H_j \circ \bar{b}_j \circ \bar{h}_j^{-1})(\bar{y}_k)|^2.$$

On choisit arbitrairement $r > 0$ et on définit les événements suivants:

$$\begin{aligned} C_{ij}^{\epsilon,r} &= C_{ij}^\epsilon \cap \{R_{ij} \geq r\}, \\ C_i^{\epsilon,r} &= C_{ij}^\epsilon \cap \{L_{ij}^\epsilon \geq 0\}, \\ C_j^{\epsilon,r} &= C_{ij}^\epsilon \cap \{L_{ij}^\epsilon < 0\}. \end{aligned}$$

Pour ϵ petit, on a

$$\begin{aligned} P(A_i \cap C_j^{\epsilon,r}) &\leq e^{-k_1/\sqrt{\epsilon}} + k_2e^{-k_3r}, \\ P(A_j \cap C_i^{\epsilon,r}) &\leq e^{-k_1/\sqrt{\epsilon}} + k_2e^{-k_3r}, \end{aligned}$$

En introduisant la fonction $\delta^{ij}_{[a,b]} : \mathbb{R}^+ \to [0,1[$,

$$\delta^{ij}_{[a,b]}(x) = P\left(\left\{|\Gamma^{-1}(H_i \circ \bar{b}_i \circ \bar{h}_i^{-1} - H_j \circ \bar{b}_j \circ \bar{h}_j^{-1})(\bar{h}_l(\hat{X}^l_t))|^2 \leq x\right\} \cap C^\varepsilon_{ij} \cap (A_i \cup A_j)\right),$$

on obtient des résultats analogues au lemme 5.11, au théorème 5.13 et au corollaire 5.16. Soit $0 < a < \tau_1 < b$, pour ε suffisamment petit,

$$P(A_i \,|\, Q^{\varepsilon,r}_j) \leq e^{-k_1/\sqrt{\varepsilon}} + \rho^\varepsilon_{[a,b]}\left(\frac{\lambda r}{\tau_1 - a}\right) k_2 e^{-k_3 r},$$

$$P(A_j \,|\, Q^{\varepsilon,r}_i) \leq e^{-k_1/\sqrt{\varepsilon}} + \rho^\varepsilon_{[a,b]}\left(\frac{\lambda r}{\tau_1 - a}\right) k_2 e^{-k_3 r},$$

où $\lambda > 1$ et $\rho^\varepsilon_{[a,b]} : \mathbb{R}^+ \to \mathbb{R}$ est une fonction croissante définie à partir de $\delta^{ij}_{[a,b]}$.

Test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy (FK_i) et (FK_j):

Soit $0 < a < d \leq b$. Pour $i < j$, on définit la statistique de test comme en (6.2) par

$$\hat{L}^\varepsilon_{ij} = \frac{1}{\varepsilon^2}\int_d^b (\hat{h}^i_t - \hat{h}^j_t) dY_t - \frac{1}{2\varepsilon^2}\int_d^b (|\hat{h}^i_t|^2 - |\hat{h}^j_t|^2)\, dt,$$

où $\hat{h}^i_t = \bar{h}_i(\hat{X}^i_t)$ et $\hat{h}^j_t = \bar{h}_j(\hat{X}^j_t)$. La partition de $\mathbb{R}^n$ étant "propre", il existe u_k un vecteur normé de $\mathbb{R}^n$ et $c_k \in \mathbb{R}^n$ tels que, par exemple,

$$A_i \subset \{\langle X_t\,,\, u_k\rangle \geq c_k\,,\, a \leq t \leq b\}.$$

Le coefficient de diffusion étant supposé constant, nous introduisons un changement de probabilité plus simple que celui proposé dans le paragraphe 6. On définit une loi de probabilité $\tilde{P}^i$ sur $(\Omega, \mathcal{F}_b)$ par

$$\frac{d\tilde{P}^i}{dP} = Z_b$$

avec

$$Z_t = \exp\left\{\int_a^{a\vee t} F^{-1}\left[(\bar{b}_i(X_t) - b(X_t)) - \frac{1}{\varepsilon}G(\bar{h}_i(X_t) - h(X_t))\right] dV_t \right.$$
$$-\frac{1}{2}\int_a^{a\vee t}\left|F^{-1}\left[(\bar{b}_i(X_t) - b(X_t)) - \frac{1}{\varepsilon}G(\bar{h}_i(X_t) - h(X_t))\right]\right|^2 dt$$
$$\left. + \frac{1}{\varepsilon}\int_a^{a\vee t}(\bar{h}_i(X_t) - h(X_t))dW_t - \frac{1}{2\varepsilon^2}\int_a^{a\vee t}\left|\bar{h}_i(X_t) - h(X_t)\right|^2 dt\right\}.$$

Les restrictions de P et $\tilde{P}^i$ à $(\Omega, \mathcal{F}_a)$ coincident, et

$$\forall D \subset A_i(a,b), \quad \tilde{P}^i(D) = P(D).$$

Pour $t \geq a$, on considère le problème de filtrage linéaire suivant:

$$\begin{cases} dX_t &= \bar{b}_i(X_t)dt + F d\tilde{V}_t^i + G d\tilde{W}_t^i, \\ dY_t &= \bar{h}_i(X_t^i)dt + \varepsilon\, d\tilde{W}_t^i, \end{cases}$$

où $\{\tilde{V}_t^i, 0 \leq t \leq b\}$ et $\{\tilde{W}_t^i; 0 \leq t \leq b\}$ sont des processus de Wiener standards indépendants sous la loi $\tilde{P}^i$. La statistique $\hat{L}_{ij}^\varepsilon$ se réécrit:

$$\hat{L}_{ij}^\varepsilon = \frac{1}{2\varepsilon^2}\int_d^b (\hat{h}_t^i - \hat{h}_t^j)^2\, dt + \frac{1}{\varepsilon}\int_d^b (\hat{h}_t^i - \hat{h}_t^j)d\tilde{\omega}_t^i$$
$$+ \frac{1}{\varepsilon^2}\int_d^b (\hat{h}_t^i - \hat{h}_t^j)(H_i\tilde{X}_t^i - H_i\hat{X}_t^i)\, dt,$$

où $\tilde{X}_t^i = \tilde{\mathbf{E}}^i(X_t \,|\, \mathcal{Y}_t)$ et $\{\tilde{\omega}_t^i; 0 \leq t \leq b\}$ est un $\mathcal{F}_t$- processus de Wiener dit processus d'innovation sous la loi $\tilde{P}^i$.

De la façon similaire à la probabilité $\tilde{P}^i$, on définit la probabilité $\tilde{P}^j$ sur $(\Omega, \mathcal{F}_b)$.

$$\hat{L}_{ij}^\varepsilon = -\frac{1}{2\varepsilon^2}\int_d^b (\hat{h}_t^i - \hat{h}_t^j)^2\, dt + \frac{1}{\varepsilon}\int_d^b (\hat{h}_t^i - \hat{h}_t^j)d\tilde{\omega}_t^j$$
$$+ \frac{1}{\varepsilon^2}\int_d^b (\hat{h}_t^i - \hat{h}_t^j)(H_j\tilde{X}_t^j - H_j\hat{X}_t^j)\, dt,$$

où $\tilde{X}_t^j = \tilde{\mathbf{E}}^j(X_t \,|\, \mathcal{Y}_t)$ et $\{\tilde{\omega}_t^j; 0 \leq t \leq b\}$ est un $\mathcal{F}_t$- processus de Wiener sous la loi $\tilde{P}^j$.

On obtient un résultat analogue à celui du corollaire 6.15.

Sous l'hypothèse ($\overline{HD}1$):

$$|H_i K_\infty^i - H_j K_\infty^j| \geq \mathcal{O}(1),$$

soit $0 < a < d < \tau_1 \leq b$, pour ε petit,

$$\begin{aligned} P(A_i \,|\, \hat{Q}_j^\varepsilon) &\leq e^{-k/\sqrt{\varepsilon}}, \\ P(A_j \,|\, \hat{Q}_i^\varepsilon) &\leq e^{-k/\sqrt{\varepsilon}}, \end{aligned}$$

avec $\hat{Q}_i^\varepsilon = C_{ij}^\varepsilon(a,b) \cap \{\hat{L}^\varepsilon(a,\tau_1) \geq 0\}$ et $\hat{Q}_j^\varepsilon = C_{ij}^\varepsilon(a,b) \cap \{\hat{L}^\varepsilon(a,\tau_1) < 0\}$.

Sous l'hypothèse ($\overline{HD}2$):

$$\begin{cases} |H_+ K_\infty^+ - H_- K_\infty^-| \leq \mathcal{O}(\varepsilon) \\ \text{et} \\ \ker\left(H_i \circ \bar{b}_i \circ \bar{h}_i^{-1} - H_j \circ \bar{b}_j \circ \bar{h}_j^{-1}\right) \subset h(\mathcal{U}), \end{cases}$$

on définit les événements tests

$$\begin{aligned} \hat{Q}_i^{\varepsilon,r} &= \hat{Q}_i^\varepsilon \cap \left\{ \frac{1}{\varepsilon^2} \int_d^b (\hat{h}_t^+ - \hat{h}_t^-)^2 ds \geq r \right\} \\ \hat{Q}_j^{\varepsilon,r} &= \hat{Q}_j^\varepsilon \cap \left\{ \frac{1}{\varepsilon^2} \int_d^b (\hat{h}_t^+ - \hat{h}_t^-)^2 ds \geq r \right\} \end{aligned}$$

Soit $\lambda > 1$, $r_0 > 0$, pour ε petit et $r \leq r_0$, on a:

$$\begin{aligned} P(A_i \,|\, \hat{Q}_j^{\varepsilon,r}) &\leq e^{-k_1/\sqrt{\varepsilon}} + \hat{\rho}_{[a,b]}^\varepsilon\left(\frac{\lambda r}{\tau_1 - d}\right) e^{-k_2 r}, \\ P(A_j \,|\, \hat{Q}_i^{\varepsilon,r}) &\leq e^{-k_1/\sqrt{\varepsilon}} + \hat{\rho}_{[a,b]}^\varepsilon\left(\frac{\lambda r}{\tau_1 - d}\right) e^{-k_2 r}, \end{aligned}$$

où la fonction $\hat{\rho}_{[a,b]}^\varepsilon$ est croissante. Elle est définie de façon analogue à (6.29) à partir de la fonction $\hat{\delta}_{[a,b]}^{ij} : \mathbb{R}^+ \to [0,1[$,

$$\begin{aligned} \hat{\delta}_{[a,b]}^{ij}(x) = P\Bigg(\Bigg\{ \inf_{[a,b]} \left| \Lambda^{-1}\left(H_i \circ \bar{b}_i \circ \bar{h}_i^{-1} - H_j \circ \bar{b}_j \circ \bar{h}_j^{-1}\right)\left(\bar{h}_i(\ddot{X}_t^i)\right) \right|^2 \leq x \Bigg\} \\ \cap C_{ij}^\varepsilon \cap (A_i \cup A_j) \Bigg). \end{aligned}$$

D'après la remarque 8.2, les intervalles de temps détectés sans passage par $\mathcal{U}$ de $t \to X_t$, sont d'autant plus courts que l est grand. Sous les hypothèses du type (2), vu que les probabilités d'erreur des tests proposés dépendent de la longueur de ces intervalles, ces tests perdent de leur efficacité lorsque l augmente.

Ainsi, pour l grand, ces procédures ne sont intéressantes que lorsque l'allure de la fonction d'observation h permet de les appliquer localement.

Bibliographie

[1] R.F. Bass, E. Pardoux : Uniqueness for diffusions with piecewise constant coefficients, in *Probability Theory Rel. Fields* Springer-Verlag, **76**, 557-572, 1987.

[2] N. El Karoui, M. Chaleyat-Maurel : Un problème de réflexion et ses applications au temps local et aux équations différentielles stochastiques sur $\mathbb{R}$, cas continu, *Astérisque* , **52-53**, 117-144, 1978.

[3] R.S. Ellis : Entropy, Large Deviations, and Statistical Mechanics, Springer-Verlag, 1985.

[4] W.H. Fleming, D. Ji, E. Pardoux : Piecewise linear filtering with smallobservation noise, in *Analysis and Optimization of Systems*, Lecture Note in Cont. and Info. Sci **111**, Springer, 725-729, 1988.

[5] W.H. Fleming, D. Ji, P. Salame et Q. Zhang : Discrete time piecewise linear filtering with small observation noise, *Brown Univ. LCDS/CCS* report **N88-27** , 1988.

[6] W.H. Fleming, E. Pardoux : Piecewise monotone filtering with small observation noise, *Siam J. Control Optim.* **20**, 261-285, 1989.

[7] N. Ikeda, S. Watanabe : Stochastic Differential equations and diffusions. North Holland, 1981.

[8] I. Karatzas, S.E. Shreve : Brownian Motion and Stochastic Calculus. Springer-Verlag, 1988.

[9] R.CH. Liptzer, A.N. Shyriaev : Statistics of Random Processes I, Springer-Verlag, 1977.

[10] E. Pardoux, M.C. Roubaud : General piecewise linear filtering problems with small observation noise, *in* 28^{th} *IEEE CDC*, Tampa, 1989.

[11] E. Pardoux, C. Savona : Piecewise linear filtering, *in Stochastic Differential Systems, Stochastic Control Theory and Appl.*, W. Fleming and P.L. Lions eds., IMA volume in Mathematics and its Appl. **10**, Springer–Verlag, 1987.

[12] J. Picard : Non linear filtering of one-dimensional diffusions in the case of a high signal -to-noise ratio, *Siam J. Apl. Math.* **46**, 1098-1125,1986.

[13] J. Picard : Asymptotic study of estimation problems with small observation noise, *in Stochastic Modelling et Filtering*, (Rome 1984), *Lect. Notes in Control and Inf. Sci.*, **91**, Springer, 1987.

[14] J. Picard : Efficiency of the extended Kalman filter for non linear systems with small noise, *Rapport INRIA* , **N1068**, 1989.

[15] W.Murray Wonham : Linear Multivariable Control: a Geometric Approach, Springer-Verlag, 1979.

Chapitre II

Filtrage linéaire par morceaux d'un système en temps discret avec petit bruit d'observation

en collaboration avec P. Milheiro de Oliveira

Table des matières

II Filtrage linéaire par morceaux d'un système en temps discret avec petit bruit d'observation **II-1**

1 Introduction II-4
2 Deux filtres de Kalman-Bucy en parallèle II-8
3 Tests de détection des passages de $\{x_k\}$ par zéro II-10
3.1 Test sur les observations y_k II-10
3.2 Test sur la sortie d'un des filtres de Kalman-Bucy II-12
4 Sous l'hypothèse (HD1): décider du signe de x_k II-19
4.1 Test de la variation quadratique II-19
4.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy II-21
5 Sous l'hypothèse ($HD2$): décider du signe de x_k II-26
5.1 Test du rapport de vraisemblance sur les acroissements des observations II-26
5.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman - Bucy II-35
6 Application II-38
6.1 Les critères de comparaison II-38
6.2 Exemples II-38
7 Conclusion II-44

A Résultats numériques et représentations graphiques des exemples présentés dans la section 6 **A-47**

1 Résultats numériques A-47
2 Représentations graphiques A-50

Résumé

Nous nous intéressons à un problème de filtrage linéaire par morceaux en temps discret avec petit bruit d'observation. Nous présentons et comparons plusieurs tests permettant de déterminer les intervalles de linéarité de la fonction d'observation, notamment dans le cas où elle est symétrique. Sur chacun de ces intervalles, nous approchons le filtre optimal par le filtre de Kalman-Bucy correspondant. Comme dans [4], nous approchons des processus discrets par des diffusions pour estimer les probabilités d'erreur et les temps moyens pour prendre une décision.

1 Introduction

On s'intéresse au problème de filtrage non linéaire suivant:

$$\begin{cases} dX_t &= b(X_t)\,dt + \sigma(X_t)\,dU_t \\ dY_t &= h(X_t)\,dt + \varepsilon\,dV_t\,, \quad Y_0 = 0\,, \end{cases} \tag{1.1}$$

où $\{X_t\,,\ t \geq 0\}$ est le processus à valeurs dans $\mathbb{R}$ non observé à estimer à l'instant τ au vu des observations jusqu'à l'instant τ du processus unidimensionnel $\{Y_t\,,\ t \geq 0\}$, ε un paramètre "petit" et $\{U_t\,,\ t \geq 0\}$ et $\{V_t\,,\ t \geq 0\}$ sont des processus de Wiener standards indépendants, à valeurs dans $\mathbb{R}$.

Il est bien connu que le problème est de dimension infinie, au sens où, pour le résoudre, on a à déterminer la solution d'une équation aux dérivées partielles, par exemple l'équation de Zakai. Si la fonction h est monotone, sous certaines hypothèses générales de "régularité", le filtre de Kalman étendu, entre autres, est une "bonne" approximation du filtre optimal (cf. [13], [8], [1], [7] et [10]). Le problème avec h non monotone, sous une certaine "hypothèse de détectabilité" a été traité dans [5].

Dans le cadre du filtrage linéaire par morceaux, on suppose que

- $b(x) = B_- x \mathbf{1}_{\{x<0\}} + B_+ x \mathbf{1}_{\{x\geq 0\}}$
- $\sigma(x) = \sigma_- \mathbf{1}_{\{x<0\}} + \sigma_+ \mathbf{1}_{\{x\geq 0\}}$
- $h(x) = H_- x \mathbf{1}_{\{x<0\}} + H_+ x \mathbf{1}_{\{x\geq 0\}}$.

Si $H_+ H_- > 0$, i.e. la fonction h est monotone, pour $\varepsilon = 0$, X_t peut être parfaitement connu et pour ε "petit", le problème n'offre pas de grandes difficultés.

Dans le cas $H_+ H_- < 0$, i.e. h non monotone, bien que $h(X_t)$ puisse être estimé de façon précise, il n'est pas immédiat qu'il en soit de même pour X_t. Le filtre de Kalman étendu est en général inefficace, le problème étant de déterminer le signe de $\{X_t\,,\ t \geq 0\}$. La détermination du signe n'étant possible que dans le cas où la variance conditionnelle est petite ([1]), on introduit l'hypothèse de "détectablité" suivante notée (HD) :

$$(HD) \qquad \begin{cases} H_-^2\sigma_-^2 \neq H_+^2\sigma_+^2 & (HD1) \\ \text{ou} & \\ H_-^2\sigma_-^2 = H_+^2\sigma_+^2 \text{ et } B_+ \neq B_- & (HD2)\,. \end{cases}$$

Sous (HD1), dans [3], un filtre approché a été proposé et dans le chapitre I, il est montré que sous (HD2) un filtre du même type peut être utilisé comme approximation du filtre optimal.

[1] Contre-exemple: $b = 0, h(x) = |x|\,, \sigma = 1$.

L'idée est de construire un filtre approché à partir de deux filtres de Kalman-Bucy, associés respectivement aux deux problèmes de filtrage linéaire suivants:

$$\begin{cases} dX_t &= B_+ X_t\, dt + \sigma_+ dU_t \\ dY_t &= H_+ X_t\, dt + \varepsilon\, dU_t \end{cases} \tag{1.2}$$

et

$$\begin{cases} dX_t &= B_- X_t\, dt + \sigma_- dU_t \\ dY_t &= H_- X_t\, dt + \varepsilon\, dU_t\,. \end{cases} \tag{1.3}$$

On calcule ces deux filtres "en parallèle" et on utilise deux tests.

- Un premier test permet de détecter des intervalles de temps durant lesquels la trajectoire de $\{X_t\}_t$ ne passe pas par 0.
- Un second test permet de décider si $X_t < 0$ ou $X_t > 0$ sur ces intervalles. sous l'hypothèse (HD1) ou (HD2).

Sur chaque intervalle de monotonie, nous approcherons le filtre optimal par le filtre de Kalman-Bucy correspondant.

Nous considérons l'intervalle de temps fini $[0, T]$.

Dans ce rapport, nous nous intéressons à la résolution numérique du problème linéaire par morceaux dans la situation où la fonction d'observation h est non monotone. Nous commençons par discrétiser le système continu (1.1) par un schéma classique de discrétisation en temps avec pas $\Delta t = \varepsilon$. Le processus $\{X_{k\Delta t}\}_k$ est approximé par $\{x_k\}_k$, $\{U_{(k+1)\Delta t} - U_{k\Delta t}\}_k$ par $\{\sqrt{\Delta t}\, u_k\}_k$, $\{V_{(k+1)\Delta t} - V_{k\Delta t}\}_k$ par $\{\sqrt{\Delta t}\, v_k\}_k$, $\{(\Delta t)^{-1}\,(Y_{(k+1)\Delta t} - Y_{k\Delta t})\}_k$ par $\{y_k\}_k$.

Nous obtenons le modèle discret suivant:

$$\begin{cases} x_{k+1} &= x_k + \varepsilon\, b(x_k) + \sqrt{\varepsilon}\, \sigma(x_k)\, u_k \\ y_k &= h(x_k) + \sqrt{\varepsilon}\, v_k\,, \end{cases} \tag{1.4}$$

avec les fonctions b, σ et h définies comme précédemment. On émet les hypothèses suivantes:

$(H1)$ $\{u_k\}_k$ et $\{v_k\}_k$ sont des bruits blancs gaussiens standards et indépendants.

$(H2)$ x_0 est une variable aléatoire réelle telle que $\mathbf{E}(\exp\{c_0 x_0^2\}) < +\infty$, pour un certain $c_0 > 0$.

$(H3)$ $H_- H_+ < 0$ et $\sigma_- \sigma_+ \neq 0$.

Sans réduire la généralité du problème, on suppose que $h(x) \geq 0\ \forall x \in \mathbb{R}$. i.e. on suppose que

$(H3')$ $H_- < 0$: $H_+ > 0$ et $\sigma_-\sigma_+ \neq 0$.

Nous adaptons l'étude faite dans le cas continu au cas discret. Notre but est de mettre en œuvre les tests permettant de séparer les intervalles de positivité et de négativité de $\{x_k\}_k$ et de comparer leur performance. Nous nous intéressons principalement à la situation où l'hypothèse $(HD2)$ est vérifiée, le problème sous l'hypothèse $(HD1)$ ayant été traité dans [4].

L'étude de l'évolution de la loi conditionnelle, obtenue de façon approchée par résolution de l'équation de Zakai discretisée (cf. [9]), nous aidera à interpréter les résultats obtenus.

Ce rapport est organisé comme il suit:

Dans le paragraphe 2, nous présenterons les deux filtres de Kalman-Bucy (2.3) et (2.4) utilisés et nous formulerons des remarques préliminaires. Dans le paragraphe 3 nous nous intéresserons à la détection des passages du signal par 0. En analogie avec le cas continu deux tests seront proposés, l'un basé sur les accroissements des observations (cf. [3]) et l'autre sur la sortie d'un des filtres de Kalman-Bucy (cf. chapitre I). Sachant que sur un intervalle de temps $[a, b]$ le signal ne passe pas par 0, avec une probabilité donnée, nous considérons un autre type de tests pour décider du signe de $\{x_k\}_k$ sur $[a, b]$. Sous l'hypothèse $(HD1)$, des études ont été faites par [4]. Ces auteurs ont mis en œuvre deux tests différents: l'un basé sur la variation quadratique, l'autre du type rapport de vraisemblance basé sur les sorties des filtres de Kalman-Bucy. Nous résumerons brièvement ces procédures dans le paragraphe 4. Sous l'hypothèse $(HD2)$, nous présenterons dans le paragraphe 5 un test de rapport de vraisemblance basé sur les accroissements des observations et adapterons le test sur les sorties des filtres de Kalman-Bucy du paragraphe précédent au cas traité.

La mise en œuvre des tests de détection et du signe ainsi décrits nécessite la détermination de formules explicites pour les bornes. Bien qu'une démonstration rigoureuse n'ait pu être faite, les formules proposées sont justifiées de manière heuristique. Dans le cas où $B_- \neq B_+$, nous émettrons l'hypothèse supplémentaire:

$(H4)$: $B_+ < 0$ et $B_- < 0$.

Cette hypothèse n'est pas trop restrictive pour notre propos. En effet, suivant le signe des coefficients de dérive, nous avons deux comportements distincts du processus $\{x_k\}$. Si B_+ et B_- sont négatifs, il devient stationnaire dès que l'intervalle de monotonie est assez grand. Sinon, il peut "fuir" vers l'infini. En fait, nous étudions la situation la plus délicate au sens où, la probabilité pour que le processus $\{x_k\}$ passe par zéro est importante. Cependant, si les valeurs $|B_-|$ et $|B_+|$ sont grandes, $\{x_k\}$ aura tendance à changer rapidement de signe et les tests n'auront probablement pas suffisament de temps pour décider.

Le paragraphe 6 est consacré à l'application numérique de ces méthodes. Des critères de comparaison entre les différents tests seront proposés et les résultats

des applications à divers exemples seront présentés. D'autre part, la solution de l'équation de Zakai sera utilisée pour la justification du comportement des filtres approchés.

Pour conclure, on présente dans le paragraphe 7 quelques réflexions sur le comportement de ces filtres et on discute la performance des différents tests étudiés.

Notation 1.1 *Etant donné un processus* $\{G_k\}$ *on écrira*

$$G_k = \mathcal{O}(\varepsilon^q),$$

où $q \in \mathbb{R}_+$, *pour signifier qu'il existe* $c_1, c_2, c_3 > 0$ *tels que*

$$E[G_k^2] \leq c_1 \exp\{-c_2 k\varepsilon\} + c_3 \varepsilon^{2q}, \forall k \in \mathcal{N}, \varepsilon > 0.$$

2 Deux filtres de Kalman–Bucy en parallèle

Soit $\mathcal{Y}_0^k$, la tribu des observations jusqu'à l'instant k, $\mathcal{Y}_0^k = \sigma\{y_0, y_1, \cdots, y_k\}$. La solution du problème de filtrage associé au système (1.1) étant donnée par la loi conditionnelle de x_k sachant $\mathcal{Y}_0^k$, nous commentons brièvement son comportement (cf. [4]). Sous l'hypothèse de "détectabilité" (HD), la variance de cette loi est "petite", sa densité se concentrant autour de deux maxima locaux, l'un négatif et l'autre positif. Ce phénomène se justifie du fait qu'on a, soit $x_k - y_k/H_- = \sqrt{\varepsilon}/H_- v_k$, soit $x_k - y_k/H_+ = \sqrt{\varepsilon}/H_+ v_k$, donc la densité conditionnelle est "petite" excepté sur deux intervalles d'amplitude d'ordre $\mathcal{O}(\sqrt{\varepsilon})$, autour de y_k/H_- et de y_k/H_+. En émettant l'hypothèse de détectabilité (HD) nous nous plaçons dans la situation où lorsque $\{x_k\}$ prend ses valeurs loin de 0 durant un certain intervalle de temps, un des pics tend à disparaître, la loi conditionnelle s'approchant d'une loi gaussienne. Il apparaît alors légitime d'approcher le filtre optimal par un filtre de Kalman–Bucy. Intuitivement on s'attend à ce que, sous l'hypothèse $(HD2)$, le temps nécessaire pour faire "disparaître" l'un des deux pics soit plus long que sous l'hypothèse $(HD1)$. Cette idée sera illustrée dans le paragraphe 6 (voir les figures A.3 à A.6 et A.12 à A.14).

Soient $(\hat{x}_k^+, Q^+)$ et $(\hat{x}_k^-, Q^-)$ les filtres de Kalman–Bucy associés respectivement aux systèmes linéaires

$$\begin{cases} x_{k+1} &= (1 + B_+ \varepsilon)\, x_k + \sqrt{\varepsilon}\, \sigma_+ u_k \\ y_k &= H_+ x_k + \sqrt{\varepsilon}\, v_k\,,\ y_0 = 0 \end{cases} \tag{2.1}$$

et

$$\begin{cases} x_{k+1} &= (1 + B_- \varepsilon)\, x_k + \sqrt{\varepsilon}\, \sigma_- u_k \\ y_k &= H_- x_k + \sqrt{\varepsilon}\, v_k\,,\ y_0 = 0\,, \end{cases} \tag{2.2}$$

avec conditions initiales gaussiennes

- $\hat{x}_0^+ = \hat{x}_0^- = E(x_0)$
- $Q_0^+ = Q^+$, $Q_0^- = Q^-$

où Q^+ et Q^- sont les variances stationnaires des lois conditionnelles.

Remarque 2.1 Les expressions de ces variances se calculent facilement. (cf. [10]).

Les filtres sont donnés par les équations suivantes:

$$\begin{cases} \hat{x}_{k+1}^+ &= (1 + \varepsilon\, B_+)\, \hat{x}_k^+ + \dfrac{1}{\varepsilon} H_+ Q^+ \left(y_{k+1} - H_+(1 + \varepsilon\, B_+)\hat{x}_k^+\right) \\ Q^- &= \varepsilon \dfrac{\rho_+ + \left[\rho_+^2 + 4\sigma_+^2 H_+^2\right]^{\frac{1}{2}}}{|H_+| \left[2 + \rho_+ + \left[\rho_+^2 + 4\sigma_+^2 H_+^2\right]^{\frac{1}{2}}\right]} \end{cases} \tag{2.3}$$

avec

$$\rho_+ = (2B_+ + B_+^2\varepsilon)\varepsilon + \sigma_+^2 H_+^2.$$

$$\left\{ \begin{array}{rcl} \hat{x}_{k+1}^- & = & (1+\varepsilon B_-)\,\hat{x}_k^- + \dfrac{1}{\varepsilon} H_- Q^- (y_{k+1} - H_-(1+\varepsilon B_-)\hat{x}_k^-) \\ Q^- & = & \varepsilon \dfrac{\rho_- + [\rho_-^2 + 4\sigma_-^2 H_-^2]^{\frac{1}{2}}}{|H_-|\left[2 + \rho_- + [\rho_-^2 + 4\sigma_-^2 H_-^2]^{\frac{1}{2}}\right]} \end{array} \right. \tag{2.4}$$

avec

$$\rho_- = (2B_- + B_-^2\varepsilon)\varepsilon + \sigma_-^2 H_-^2.$$

Remarque 2.2 Le fait d'avoir considéré une condition initiale gaussienne avec pour variance, la variance conditionnelle stationnaire, n'est pas gênant pour la suite de notre propos. En effet, puisque les systèmes étudiés se caractérisent par une "mémoire courte", l'influence de la condition initiale tend à disparaître rapidement.

Les gains stationnaires sont donc,

$$K^+ = \frac{1}{\varepsilon} H_+ Q^+ \qquad \text{et} \qquad K^- = \frac{1}{\varepsilon} H_- Q^- .$$

On introduit les processus dits "d'innovation"

$$\nu_{k+1}^+ = y_{k+1} - H_+(1+\varepsilon B_+)\hat{x}_k^+$$

et

$$\nu_{k+1}^- = y_{k+1} - H_-(1+\varepsilon B_-)\hat{x}_k^- .$$

Les processus ainsi définis sont considérés approximativement comme des "bruits blancs" au sens où leurs variances sont d'ordre $\mathcal{O}(\varepsilon)$, tandis que les corrélations $E(\nu_{k+1}^+ \nu_k^+)$ et $E(\nu_{k+1}^- \nu_k^-)$ sont d'ordre $\mathcal{O}(\varepsilon^2)$.

On notera $\varepsilon\vartheta^+$ et $\varepsilon\vartheta^-$ leurs variances respectives,

$$\begin{aligned} \varepsilon\vartheta^+ &= \varepsilon(1+\sigma_+^2 H_+^2) + H_+^2(1+\varepsilon B_+)^2 Q^+, \\ \varepsilon\vartheta^- &= \varepsilon(1+\sigma_-^2 H_-^2) + H_-^2(1+\varepsilon B_-)^2 Q^- . \end{aligned}$$

Dans la suite, on étudiera la procédure de test permettant de séparer les intervalles de monotonie de la fonction h et de décider du signe de x_k. Sur chacun de ces intervalles, il sera alors possible d'approcher le filtre optimal par le filtre de Kalman-Bucy correspondant.

3 Tests de détection des passages de $\{x_k\}$ par zéro

Nous devons premièrement déterminer sur l'intervalle de temps $[0, T]$, les intervalles sur lesquels $\{x_k\}$ ne passe pas par zéro, avec une probabilité donnée , proche de 1. Nous présentons deux tests de détection, le premier s'appliquant aux observations y_k, le second à la sortie de l'un des filtres de Kalman-Bucy.

Soit l'intervalle de temps $[a, b]$, avec $0 < a < b < T$. On pose $m = [(b-a)/\varepsilon]$ et $i_0 = [a/\varepsilon]$. On considère les deux événements suivants:

$$\begin{aligned} A_- &= \{x_k < 0\,;\, k = i_0, i_0+1, \cdots, i_0+m\}, \\ A_+ &= \{x_k > 0\,;\, k = i_0, i_0+1, \cdots, i_0+m\}. \end{aligned}$$

On remarque que

$$(A_- \cup A_+)^c = \{\, x_k\, x_{k+1} < 0 \text{ pour un certain k }, i_0 \le k \le i_0+m\}.$$

3.1 Test sur les observations y_k

Etant donné que "$h(x) = 0$ ssi $x = 0$", $h(x_k)$ "petit" implique par continuité de l'application h, x_k proche de zéro. Cependant $h(x_k)$ n'est pas observé, mais pour ε "petit", il est approché par y_k.

Soit $c_{obs} > 0$ une constante à déterminer.

On définit l'événement test suivant:

$$\mathcal{C} = \{|y_k| \ge c_{obs}\}.$$

On montre de façon similaire à la version en temps continu (cf. [3][proposition 3.1]) la proposition suivante (cf. [4][proposition 2.1]):

Théorème 3.1 *Soit $c_{obs} > 0$ et $\varepsilon_0 > 0$ donnés, il existe $\beta > 0$ tel que $\forall \varepsilon \in (0.\varepsilon_0]$ on a*

$$P(\{(A_- \cup A_+)^c | \mathcal{C}\}) \le e^{-\beta/\varepsilon}.$$

Détermination de la constante c_{obs}:

Nous utilisons, pour la détermination de la constante c_{obs}, un raisonnement similaire à celui proposé dans [4], sauf que dans notre cas le coefficient de dérive n'est plus supposé constant mais constant par morceaux. Intuitivement le choix de c_{obs} doit être un compromis entre "conserver des intervalles de monotonie suffisament grands" et "réduire la probabilité d'erreur du test".

On cherche à ce que la probabilité

$$P(\{x_k x_{k+1} < 0\} \mid \{|y_k| \geq c_{obs}, |y_{k+1}| \geq c_{obs}\})$$

soit faible. D'après l'hypothèse $(H3)$ on obtient:

$$P(\{y_k < -c_{obs}\}) \leq P(\{v_k < -c_{obs}/\sqrt{\varepsilon}\}) \leq \frac{1}{2} e^{-c_{obs}^2/(2\varepsilon)}.$$

Nous nous permettons donc de considérer uniquement le conditionnement par rapport à l'événement $\{y_k \geq c_{obs}, y_{k+1} \geq c_{obs}\}$. De plus, on a:

$$\{x_k x_{k+1} < 0\} \cap \{y_k \geq c_{obs}, y_{k+1} \geq c_{obs}\} = D^1_{c_{obs}} \cup D^2_{c_{obs}},$$

avec les définitions suivantes pour les évènements D^1_c, D^2_c:

$$\begin{aligned} D^1_c &= \{x_k < 0, x_{k+1} > 0\} \cap \{y_k \geq c, y_{k+1} \geq c\}, \\ D^2_c &= \{x_k > 0, x_{k+1} < 0\} \cap \{y_k \geq c, y_{k+1} \geq c\}, \end{aligned}$$

pour tout $c > 0$.

On considère séparemment les évènements $D^1_{c_1}$ et $D^2_{c_2}$.

- *Sur* $D^1_{c_1}$, on a les récurrences suivantes:

$$\begin{aligned} x_{k+1} &= (1 + \varepsilon B_-)x_k + \sigma_- \sqrt{\varepsilon}\, u_k, \\ y_k &= H_- x_k + \sqrt{\varepsilon}\, v_k, \quad y_{k+1} = H_+ x_{k+1} + \sqrt{\varepsilon}\, v_{k+1}. \end{aligned}$$

On déduit l'égalité:

$$\frac{1}{\sqrt{\varepsilon}}(H_- y_{k+1} - H_+(1 + \varepsilon B_-)y_k) = H_+ H_- \sigma_- u_k + H_- v_{k+1} - H_+(1 + \varepsilon B_-)v_k.$$

Or, d'après $(H3')$ on a $h(x_k) \geq 0$ et, pour ε suffisamment petit, on obtient la majoration

$$\frac{1}{\sqrt{\varepsilon}} |H_- - H_+(1 + \varepsilon B_-)|\, c_1 \leq |Z_k|,$$

où

$$Z_k = H_+ H_- \sigma_- u_k + H_- v_{k+1} - H_+(1 + \varepsilon B_-)\, v_k$$

suit une loi gaussienne $\mathcal{N}(0, H_+^2 H_-^2 \sigma_-^2 + H_-^2 + H_+^2(1 + \varepsilon B_-)^2)$.

Soit $\{Z_k^0\}$ le processus normalisé associé à $\{Z_k\}$. Pour un risque α donné [2], on détermine $\lambda > 0$ tel que

$$P(|Z_k^0| \geq \lambda) \leq \alpha,$$

par inversion de la fonction de répartition de la loi normale centrée réduite. On obtient ainsi l'expression de la constante c_1:

$$c_1 = \lambda \sqrt{\varepsilon} \frac{\sqrt{H_+^2 H_-^2 \sigma_-^2 + H_-^2 + H_+^2(1 + \varepsilon B_-)^2}}{|H_- - H_+(1 + \varepsilon B_-)|}.$$

[2] Généralement on prend $\alpha = 0.05$.

• *Sur* $D^2_{c_2}$, un raisonnement similaire au précédent nous amène à considérer une constante c_2 donnée par:

$$c_2 = \lambda\sqrt{\varepsilon}\frac{\sqrt{H_+^2 H_-^2\sigma_+^2 + H_+^2 + H_-^2(1+\varepsilon B_+)^2}}{|H_+ - H_-(1+\varepsilon B_+)|}.$$

Dans les applications numériques, pour une probabilité d'erreur α donnée, nous prendrons $c_{obs} = \max(c_1, c_2)$.

Remarque 3.2 Dans les calculs de la constante c_{obs}, la prise en compte de la dérive introduit uniquement des termes d'ordre $\mathcal{O}(\varepsilon^{3/2})$ que nous négligerons devant les termes d'ordre $\mathcal{O}(\sqrt{\varepsilon})$.

3.2 Test sur la sortie d'un des filtres de Kalman–Bucy

On se propose de décrire un test capable de détecter les changements de signe éventuels de x_k à l'aide de la sortie obtenue par un des deux filtres de Kalman–Bucy, par exemple $\{\hat{x}_k^+\}$. La construction d'un tel test se justifie par le fait que "$h(x) = 0$ ssi $x = 0$" et par la proposition suivante:

Proposition 3.3 *Soit $0 < a < b$. Pour tout $\theta > 0$, il existe $\beta > 0$, $\varepsilon_0 > 0$ tels que, pour $\varepsilon \in (0, \varepsilon_0]$, si on définit $i_0 = [a/\varepsilon]$ et $m = [(b-a)/\varepsilon]$, on a:*

$$P(\{\max_{k=i_0,\cdots,i_0+m} |h(x_k) - H_+\hat{x}_k^+| > \theta\}) \leq e^{-\beta/\sqrt{\varepsilon}}.$$

En vu de la démonstration de cette proposition, on établit d'abord le lemme suivant:

Lemme 3.4 *Pour tout $T > 0$, il existe $\varepsilon_0 > 0$ et $c, C > 0$ tels que, pour tout $\varepsilon \in (0, \varepsilon_0]$, $m = [T/\varepsilon]$,*

$$\mathbf{E}(\exp \max_{k=0,\cdots,m} c\,x_k^2) \leq C.$$

Notation 3.5 *On notera c, une constante indépendante de ε, sans se soucier de sa valeur. Elle pourra donc varier d'une ligne à l'autre.*

Preuve

La preuve de ce lemme est présentée en 3 étapes (cf. chapitre I [lemme 2.7] pour la preuve analogue en temps continu).

Etape 1:

On démontre le résultat suivant:

Lemme 3.6 *On définit le processus* $\{z_k\}$ *par*

$$z_{k+1} = (1 + \varepsilon B)z_k + \sqrt{\varepsilon}\,\sigma u_k\,, \tag{3.1}$$

où z_0 *est une variable aléatoire indépendante du processus* $\{u_k\}$ *et il existe des constantes* $c_0\,,\ C_0 > 0$ *telles que*

$$\mathbf{E}(\exp c_0\, z_0^2) < C_0\,. \tag{3.2}$$

Alors, pour tout $T > 0$, *il existe* $\varepsilon_0 > 0$ *et* $c\,,\ C > 0$ *tels que, pour tout* $\varepsilon \in (0, \varepsilon_0]$, $m = [T/\varepsilon]$,

$$\mathbf{E}(\exp \max_{k=0,\cdots,m} c\, x_k^2) \leq C\,.$$

Preuve

Nous réécrivons l'expression de z_{k+1} en mettant en évidence une partie dépendante de z_0 et une partie martingale:

$$z_{k+1} = (1 + \varepsilon B)^k[(1 + \varepsilon B)z_0 + M_k]\,,$$

avec

$$M_k = \sigma\sqrt{\varepsilon} \sum_{j=0}^{k} (1 + \varepsilon B)^{-j} u_j\,.$$

$$z_{k+1}^2 \leq (1 + \varepsilon B)^{2k}[2(1 + \varepsilon B)^2 z_0^2 + 2M_k^2] \leq e^{2TB}[2(1 + \varepsilon B)^2 z_0^2 + 2M_k^2]$$

D'après (3.2) et l'inégalité de Cauchy-Schwartz, il est suffisant de montrer qu'il existe $\varepsilon_0\,,\ c\,,\ C > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\mathbf{E}(\exp \max_{k=0,\cdots,m-1} c\, M_k^2) \leq C\,.$$

Or $\{M_k\}$ est un processus gaussien centré de variance $\mathbf{E}[M_k^2]$,

$$\mathbf{E}[M_k^2] = k\varepsilon\sigma^2 \qquad \text{si } B = 0,$$

$$\mathbf{E}[M_k^2] = \sigma^2\, \frac{(1 + \varepsilon B)^2 - (1 + \varepsilon B)^{-2k}}{B(2 + \varepsilon B)} \qquad \text{si } B \neq 0.$$

Si $B > 0$, $\qquad \mathbf{E}[M_k^2] \leq \sigma^2 \dfrac{(1 + \varepsilon B)^2}{B(2 + \varepsilon B)}$

$$\text{Si } B<0, \quad \mathbf{E}[M_k^2] \le \sigma^2 \frac{-(1+\varepsilon B)^2 + \exp(2k\varepsilon B + k\varepsilon^2 B^2)}{-B(2+\varepsilon B)}$$

Il existe donc ε_0 et une constante $C = C(\varepsilon_0, T) > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\max_{k=0,m-1} \mathbf{E}[M_k^2] \le C\,.$$

On en déduit l'existence de $c_1 > 0$ tel que

$$\limsup_{\varepsilon \to 0} \mathbf{E}(\exp c_1 M_{m-1}^2) \le C\,.$$

Vu que $\{\exp c M_k^2\}$ est une sous-martingale, on a:

$$\mathbf{E}(\max_{k=0,\cdots,m-1} \exp c M_k^2) \le 4\,\mathbf{E}(\exp c_1 M_{m-1}^2) \le C\,,$$

pour $c \le c_1/2$, d'où le lemme.

□

Etape 2:

On introduit les notations:

$$B = \sup(|B_+|, |B_-|)$$

$$\sigma = \sup(|\sigma_+|, |\sigma_-|)$$

On démontre que les moments d'ordre $2p$, $p \in \mathcal{N}$, du processus $\{x_k\}$ peuvent être majorés par les moments d'ordre $2p$ du processus $\{z_k\}$ défini par

$$\begin{cases} z_{k+1} &= (1+\varepsilon B) z_k + \sqrt{\varepsilon}\,\sigma u_k\,, \\ z_0 &= x_0\,. \end{cases}$$

En effet, d'après l'équation d'état, on a

$$x_{k+1}^{2p} = \sum_{j=0}^{2p} C_j^{2p} [x_k + \varepsilon\, b(x_k)]^j [\sqrt{\varepsilon}\,\sigma(x_k)\, u_k]^{2p-j}\,.$$

Le fait que $\{u_k\}$ est un bruit blanc et que u_k est indépendant de x_k, entraine l'égalité

$$\mathbf{E}[x_{k+1}^{2p}] = \sum_{j=0;\,j \text{ pair}}^{2p} C_j^{2p}\, \mathbf{E}[(x_k + \varepsilon\, b(x_k))^j (\sqrt{\varepsilon}\,\sigma(x_k))^{2p-j}] \mathbf{E}[u_k^{2p-j}]\,.$$

On obtient la majoration suivante:

$$\mathbf{E}[x_{k+1}^{2p}] \leq \sum_{j=0;j \text{ pair}}^{2p} C_j^{2p} (\sqrt{\varepsilon}\sigma)^{2p-j} \frac{(2p-j)!}{2^{p-j/2}(p-j/2!)} (1+\varepsilon B)^j \mathbf{E}[x_k^j] .$$

D'autre part, on note que

$$\mathbf{E}[z_{k+1}^{2p}] = \sum_{j=0;j \text{ pair}}^{2p} C_j^{2p} (\sqrt{\varepsilon}\sigma)^{2p-j} \frac{(2p-j)!}{2^{p-j/2}(p-j/2)!} (1+\varepsilon B)^j \mathbf{E}[z_k^j] .$$

Vu que $z_0 = x_0$, on obtient

$$\mathbf{E}[x_k^{2p}] \leq \mathbf{E}[z_k^{2p}] , \; k = 0, 1, \cdots, [T/\varepsilon] .$$

Le lemme 3.6 entraine l'existence de $\varepsilon_0 > 0$ et c, $C > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\mathbf{E}(\exp \max_{k=0,\cdots,m} c x_k^2) \leq C . \tag{3.3}$$

Etape 3:

Nous utilisons la majoration (3.3) pour obtenir le lemme 3.4. D'après la formule de Taylor, on a:

$$\begin{aligned} \exp c x_{j+1}^2 &= \exp c x_j^2 + 2c x_j \exp(c x_j^2)(x_{j+1} - x_j) \\ &\quad + 2c \exp(c\theta_j^2)(1 + 2c\theta_j^2)(x_{j+1} - x_j)^2 . \end{aligned}$$

On obtient l'expression

$$\begin{aligned} \exp c x_{k+1}^2 &= \exp c x_0^2 + 2c \sum_{j=0}^{k} x_{k-j} \exp(c x_{k-j}^2)(x_{k+1-j} - x_{k-j}) \\ &\quad + 2c \sum_{j=0}^{k} \exp(c\theta_{k-j}^2)(1 + 2c\theta_{k-j}^2)(x_{k+1-j} - x_{k-j})^2 . \end{aligned}$$

Puisque $|\theta_k| \leq |x_k| + |x_{k+1}|$, on a la majoration

$$\begin{aligned} & \mathbf{E}[\max_{k=0,\cdots,m-1} \exp c x_{k+1}^2] \\ \leq \; & \mathbf{E}[\exp c x_0^2] + 2c\varepsilon B \sum_{j=0}^{m-1} \mathbf{E}[x_{k-j}^2 \exp(c x_{k-j}^2)] \\ & + 2c\mathbf{E}\left[\max_{k=0,\cdots,m-1} \sqrt{\varepsilon} \left(\sum_{j=0}^{k} \exp(c\theta_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j} \right) \right] \\ & + 4c\varepsilon \sum_{j=0}^{m-1} \mathbf{E}\left[\exp\left(2c(x_{k-j}^2 + x_{k+1-j}^2) \right) \left(1 + 4c(x_{k+1-j}^2 + x_{k-j}^2) \right) \right. \\ & \qquad \left. (\varepsilon B^2 x_{k-j}^2 + \sigma^2 u_{k-j}^2) \right] . \end{aligned} \tag{3.4}$$

En utilisant l'hypothèse ($H2$), l'inégalité de Cauchy-Schwartz et le résultat (3.3) nous remarquons que le lemme est démontrée s'il existe $\varepsilon_0 > 0$ et c, $C > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\mathbf{E}[\max_{k=0,\cdots,m-1} \sqrt{\varepsilon} \sum_{j=0}^{k} \exp(c, x_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j}] \leq C .$$

Le processus

$$M_k = \sqrt{\varepsilon} \sum_{j=0}^{k} \exp(c\, x_{k-j}^2) x_{k-j}\, \sigma(x_{k-j})\, u_{k-j}$$

est une $\mathcal{F}_k$-martingale avec $\mathcal{F}_k = \sigma(x_0, u_0, \cdots, u_k)$.
On utilise une propriété des martingales:

$$\mathbf{E}[\max_{k=0,\cdots,m-1} M_k^2] \leq 4\, \mathbf{E}[M_{m-1}^2] .$$

Vu que u_k et x_k sont indépendants et x_k est $\mathcal{F}_{k-1}$ - mesurable, on a donc

$$\begin{aligned} \mathbf{E}[M_{m-1}^2] &\leq (m-1)\, \varepsilon\, \sigma \max_{j=0,\cdots,m-1} \mathbf{E}[\exp(4\, c\, x_{m-1-j}^2)]^{1/2} \mathbf{E}[x_{m-1-j}^4]^{1/2} \\ &\leq Tc, \end{aligned}$$

d'après l'inégalité (3.3), d'où le lemme 3.4.

□

Preuve (de la proposition 3.3)

Nous utilisons la notion de sous-différentiel de h.

Puisque la fonction h est convexe, il existe ρ_k appartenant au sous-différentiel de h à un point θ_k tel que

$$h(x_{k+1}) = h(x_k) + \rho_k(x_{k+1} - x_k)$$

et

$$|\rho_k| \leq \sup(|H_+|, |H_-|) \triangleq H .$$

On définit

$$Z_k \triangleq h(x_k) - H_+ \hat{x}_k^+ .$$

D'après l'équation du filtre $\hat{x}_k^+$ et l'équation d'état, on obtient l'expression récurrente:

$$Z_{k+1} = \delta Z_k + \varepsilon\, b(x_k)\rho_k(1 - H_+K_+) + \sqrt{\varepsilon}\, \sigma(x_k)\rho_k(1 - H_+K_+)u_k + \sqrt{\varepsilon}\, H_+K_+v_{k+1} .$$

où

$$\delta \triangleq (1 + \varepsilon B_+)(1 - H_+K_+) .$$

On remarque que $0 < \delta < 1$. On utilise cette récurrence jusqu'au terme Z_{i_0}. On établit ainsi la majoration:

$$\begin{aligned} |Z_{k+1}| \;\leq\; & \exp\{-c(k+1)\}|Z_{i_0}| + c\varepsilon \sum_{j=0}^{k-i_0} \delta^j |x_{k-j}| \\ & + c\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}| + c\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |v_{k+1-j}| \,. \end{aligned} \tag{3.5}$$

On traite les quatre termes séparemment. L'inégalité de Bienaymé Tchebichev et le fait que $\mathbf{E}[|Z_{i_0}|] < c$ avec $i_0 = \left[\frac{a}{\varepsilon}\right]$ implique

$$P\left(|Z_{i_0}| > \frac{\theta}{4} e^{ci_0}\right) \leq \frac{c}{\theta} e^{-ca/\varepsilon} \,.$$

D'autre part, en utilisant le lemme 3.4, on obtient la majoration

$$P\left(\varepsilon \sum_{j=0}^{k-i_0} \delta^j |x_{k-j}| > \frac{\theta}{4}\right) \leq c e^{-\theta(1-\delta)/(4\varepsilon)}$$

Le lemme 2.4 dans le chapitre I entraine

$$\begin{aligned} & P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}|\right) \\ \leq\; & P\left(\max_{k=i_0,\cdots,i_0+m-1} |u_k| > \frac{\theta(1-\delta)}{4c\sqrt{\varepsilon}}\right) \leq \frac{b-a}{\varepsilon} e^{-c\theta/\varepsilon} \,, \end{aligned}$$

donc il existe $\varepsilon_0 > 0$ tel que $\forall \varepsilon \in (0, \varepsilon_0]$.

$$P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}|\right) \leq e^{-c/\varepsilon} \,.$$

De même pour le 4^{me} terme à droite dans l'expression (3.5), on a:

$$P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |v_{k-j}|\right) \leq e^{-c/\varepsilon} \,.$$

De l'expression (3.5), en utilisant les quatre majorations ci-dessus, on obtient la proposition 3.3.

□

Soit c_{FK} une constante strictement positive à déterminer.
On définit l'évènement test suivant:

$$\hat{\mathcal{C}} = \{|\hat{x}_k^+| \geq c_{FK} \,;\, k = i_0, i_0+1, \cdots, i_0+m\} \,.$$

Comme conséquence de la proposition 3.3, on obtient le théorème suivant:

Théorème 3.7 *Il existe $\beta > 0$, $\varepsilon_0 > 0$ tels que, pour $\varepsilon \in (0, \varepsilon_0]$, on a:*

$$P((A_+ \cup A_-)^c \,|\hat{C}) \leq e^{-\beta/\sqrt{\varepsilon}}.$$

Preuve La preuve de ce théorème suit pas à pas celle du résultat équivalent en temps continu (cf. chapitre I [Théorème 3.4]).

□

Détermination de la constante c_{FK}:

On cherche une constante c_{FK} telle que la probabilité

$$P\left(\{x_k x_{k+1} \leq 0\} \,|\, \{|\hat{x}_k^+| \geq c_{FK}\,,\, |\hat{x}_{k+1}^+| \geq c_{FK}\}\right)$$

soit faible.

On considère y_k comme une approximation de $h(x_k)$. Dans l'étude précédente, pour une probabilité d'erreur α donnée, on a déterminé une constante c_{obs} permettant de tester au vu de y_k, l'occurence d'un changement du signe de x_k. Or, un changement de signe de x_k est directement lié au fait que $h(x_k)$ devient "petit". Si $H_+\hat{x}_k^+$ était une meilleure approximation de $h(x_k)$ que y_k, tout au moins aux alentours de zéro, la probabilité d'erreur du test "$\hat{x}_k^+ \geq c_{obs}/H_+$" serait minorée par α. Nous ne savons pas si cette condition est vérifiée, mais, sur l'évènement A_+, le filtre $\hat{x}_m^+$ est une meilleure approximation de x_m que y_m/H_+. Et, en tout cas, la proposition 3.3 nous assure que l'écart $|h(x_k) - H_+\hat{x}_k^+|$ est petit avec une probabilité proche de 1. Il semble alors raisonnable de penser que le test "$\hat{x}_k^+ \geq c_{obs}/H_+$" permettra de détecter convenablement les passages de x_k par 0. Nous proposons donc pour constante

$$c_{FK} = c_{obs}/\,|H_+|\,.$$

Le même type de raisonement est valable quand le test est appliqué sur le filtre $\hat{x}_k^-$ au lieu du filtre $\hat{x}_k^+$. Nous prenons lors

$$c_{FK} = c_{obs}/\,|H_-|\,.$$

Sans perte de généralité on considèrera par la suite un seul intervalle $[a, b]$ représentant l'intervalle de temps durant lequel il n'y a pas de passage à 0, avec une probabilité d'erreur donnée, i_0 et m seront définis comme en début de cette section.

4 Sous l'hypothèse (HD1): décider du signe de x_k

Dans ce paragraphe, nous présentons deux tests permettant de choisir entre les évènements A_+ et A_-, dans le cas où le problème de filtrage linéaire par morceaux vérifie l'hypothèse $(HD1)$. Ces tests ont été etudiés dans [4]. Le problème analogue en temps continu a éte traité dans [3]. Le premier test considéré, dit de la variation quadratique, n'est applicable que sous l'hypothèse $(HD1)$. Le second test, du type rapport de vraisemblance sur les sorties du filtre de Kalman–Bucy, pourra être adapté au cas où l'hypothèse $(HD2)$ est vérifiée (cf. paragraphe 5.2). Le premier test sera donc introduit de façon brève. Le deuxième test sera décrit de façon plus détaillée, puisque la même idée sera utilisée par la suite.

Dans [4], les auteurs ont mis en œuvre ces tests sous deux formes: test à taille d'échantillon fixée et test séquentiel. D'après leur étude, les tests séquentiels semblent plus intéressants que ceux à taille d'échantillon fixée, vis a vis du critère du "temps moyen pour prendre une décision". Nous ne présenterons donc que les tests séquentiels.

4.1 Test de la variation quadratique

On considère un test d'hypothèses basé sur la variation quadratique de la suite d'observations y_k pour décider entre les deux alternatives "A_+" et "A_-". Cette décision est prise sur N observations ($0 < N \leq m$) dans un intervalle de monotonie obtenu par un des tests introduits dans le paragraphe 3. On considère N comme étant un temps d'arrêt.
Sans restreindre la généralité du problème, on suppose que $B_+ = B_- = B$.

Remarque 4.1 Le fait de considérer B_+ et B_- distincts a pour unique conséquence l'introduction de termes négligeables dans les expressions, la procédure restant essentiellement la même.

Notons:

$$\begin{aligned} \Delta_k &= y_{k+1} - (1 + B\varepsilon)\, y_k\,, \\ \Upsilon_+^2 &= H_+^2\sigma_+^2 + 1 + (1 + B\varepsilon)^2\,, \\ \Upsilon_-^2 &= H_-^2\sigma_-^2 + 1 + (1 + B\varepsilon)^2. \end{aligned}$$

Alors

- si $x_k \geq 0$ et $x_{k+1} \geq 0$, la variable aléatoire Δ_k suit une loi gaussienne de moyenne nulle et de variance $\Upsilon_+^2\varepsilon$
- si $x_k < 0$ et $x_{k+1} < 0$, la variable aléatoire Δ_k suit une loi gaussienne de moyenne nulle et de variance $\Upsilon_-^2\varepsilon$.

L'hypothèse de détectabilité ($HD1$) implique

$$\Upsilon_+^2 \neq \Upsilon_-^2 .$$

Le problème de décider entre les alternatives "A_+" et "A_-" conduit à un test d'hypothèses sur ces variances.
On définit le rapport logarithmique

$$Z_k = \ln \frac{\mathcal{N}(\Delta_k, 0, \Upsilon_+^2 \varepsilon)}{\mathcal{N}(\Delta_k, 0, \Upsilon_-^2 \varepsilon)},$$

où $\mathcal{N}(x, \mu, \gamma^2)$ représente la densité de la loi gaussienne de paramètres μ et γ^2 et la série

$$S_n = \sum_{k=i_0}^{i_0+n} Z_k ,$$

où $0 \leq n \leq m$, i_0 et m étant définis dans le paragraphe 3.
On obtient les expressions suivantes:

$$Z_k = \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2}\left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2}\right)\frac{\Delta_k^2}{\varepsilon}$$

et

$$S_n = (n+1)\log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2}\left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2}\right)\sum_{k=i_0}^{i_0+n} \frac{\Delta_k^2}{\varepsilon}.$$

Soit $l_1 > 0$ et $l_2 > 0$ fixés, on introduit le temps d'arrêt N^*:

$$N^* = \inf\{n > 0 : S_n < -l_1 \text{ ou } S_n > l_2\} \wedge m .$$

On décrit le test comme suit: si $S_{N^*} \geq l_2$ on accepte "A_+" et si $S_{N^*} \leq -l_1$ on accepte "A_-"; sinon on dira que le test ne permet pas de décider.

Il reste à proposer une méthode de détermination des bornes l_1 et l_2. On approche le processus discret $\{\varepsilon S_n\}$ par un certain processus de diffusion $\{\zeta_t\}$ dont on écrit les équations sous "A_+" et sous "A_-". On obtient ainsi, par la formule de Dynkin, des équations différentielles ordinaires permettant de calculer les temps moyens pour prendre une décision $\mathbf{E}(T_+^*)$ et $\mathbf{E}^-(T_-^*)$ et pour les probabilités d'erreur p_+ et p_- (cf. [Fleming-Rishel]).
On introduit les notations

$$\mu_+ = \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2}\left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2}\right)\Upsilon_+^2$$

$$\mu_- = \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2}\left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2}\right)\Upsilon_-^2$$

$$\gamma_+ = \frac{1}{2}\left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2}\right)\sqrt{\Upsilon_+^4 + 2(1+B\epsilon)^2}$$

$$\gamma_- = \frac{1}{2}\left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2}\right)\sqrt{\Upsilon_-^4 + 2(1+B\epsilon)^2}.$$

La résolution des équations nous donne comme solution:

$$p_+ = \frac{1 - e^{-\mu_+/\gamma_+^2 l_2}}{e^{\mu_+/\gamma_+^2 l_1} - e^{-\mu_+/\gamma_+^2 l_2}} \tag{4.1}$$

$$p_- = \frac{1 - e^{\mu_-/\gamma_-^2 l_1}}{e^{-\mu_-/\gamma_-^2 l_2} - e^{\mu_-/\gamma_-^2 l_1}} \tag{4.2}$$

$$\mathbf{E}(T_+^*) = \left(\frac{l_2}{\mu_+} - p_+\frac{l_2+l_1}{\mu_+}\right)\epsilon$$

$$\mathbf{E}(T_-^*) = \left(-\frac{l_1}{\mu_-} + p_-\frac{l_2+l_1}{\mu_-}\right)\epsilon,$$

où

- p_+ est "la probabilité de refuser "A_+" à tort",
- p_- est "la probabilité de refuser "A_-" à tort",
- T_+^* et T_-^* sont les temps d'arrêt

$$T_+^* = \inf\{t : \zeta_t \geq \epsilon\, l_2 \text{ ou } \zeta_t \leq -\epsilon\, l_1\}, \text{ dans le cas positif,}$$

et

$$T_-^* = \inf\{t : \zeta_t \geq \epsilon\, l_2 \text{ ou } \zeta_t \leq -\epsilon\, l_1\}, \text{ dans le cas négatif .}$$

Pour des probabilités d'erreur p_+ et p_- données ([3]), les expressions (4.1) et (4.2) permettent de calculer les bornes l_1 et l_2.

4.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman–Bucy

On décrit un autre test du type rapport de vraisemblance pour décider entre les deux alternatives "A_+" et "A_-". L'idée de ce test est la suivante:

[3]Généralement on prend $p_+ = p_- = 0.05$.

Soit i_1 un entier, $i_0 < i_1 < i_0 + m$. Intuitivement, on s'attend à ce que $\hat{h}_k = E(h(X_k)|\mathcal{Y}_0^k)$ soit très proche de $H_+\hat{x}_k^+$ sur A_+ et de $H_-\hat{x}_k^-$ sur A_-, pour $i_1 \leq k \leq i_0 + m$. On a

$$\begin{aligned} \text{sur } A_+ : \quad y_k &= H_+(1+\epsilon B_+)\hat{x}_{k-1}^+ + \nu_k^+, \\ \text{sur } A_- : \quad y_k &= H_-(1+\epsilon B_-)\hat{x}_{k-1}^- + \nu_k^-, \end{aligned}$$

où $\{\nu_k^+\}$ et $\{\nu_k^-\}$ peuvent être considérés comme des bruits blancs gaussiens dits "processus d'innovation" (cf. section 2).

On pose

$$Z_k = H_+(1+\epsilon B_+)\hat{x}_k^+ - H_-(1+\epsilon B_-)\hat{x}_k^-$$

et on prend pour statistique de test le rapport logarithmique suivant:

$$\hat{L}_n = \frac{1}{\epsilon}\left[\sum_{k=i_1}^{i_1+n} Z_k\, y_{k+1} - \frac{1}{2}\sum_{k=i_1}^{i_1+n}\left((H_+(1+\epsilon B_+)\hat{x}_k^+)^2 - (H_-(1+\epsilon B_-)\hat{x}_k^-)^2\right)\right],$$

avec $i_1 \leq n \leq m$.

La statistique $\hat{L}_n$ se réécrit sous la forme suivante:

$$\begin{aligned} \text{Sur } A_+ : \hat{L}_n &= +\frac{1}{2\epsilon}\sum_{k=i_1}^{i_1+n} Z_k^2 + \frac{1}{\epsilon}\sum_{k=i_1}^{i_1+n} Z_k \nu_{k+1}^+ . \\ \text{Sur } A_- : \hat{L}_n &= -\frac{1}{2\epsilon}\sum_{k=i_1}^{i_1+n} Z_k^2 + \frac{1}{\epsilon}\sum_{k=i_1}^{i_1+n} Z_k \nu_{k+1}^- . \end{aligned}$$

On note $\hat{R}_n$ le terme de précision du test:

$$\hat{R}_n = \frac{1}{\epsilon}\sum_{k=i_1}^{i_1+n} Z_k^2 .$$

Pour R_n "suffisamment grand", le signe de $\hat{L}_n$ permet de choisir entre les alternatives "A_-" et "A_+".

Soit $l_1 > 0$ et $l_2 > 0$ des bornes fixées, on introduit le temps d'arrêt N^*:

$$N^* = \inf\{n \geq i_1 \,:\, \hat{L}_n \geq l_2 \text{ ou } \hat{L}_n \leq -l_1\} \wedge m .$$

On applique la règle de décision qui suit.

Règle de décision:

- Si $\hat{L}_{N^*} \geq l_2$, on décide "A_+".

- Si $\hat{L}_{N^\bullet} \leq -l_1$, on décide "$A_-$".

- Sinon, on ne peut pas décider.

Il nous reste à calculer les bornes l_1 et l_2 pour une probabilité d'erreur donnée α et à estimer les temps moyens pour atteindre une décision. Le raisonnement présenté par la suite constitue une justification purement heuristique du choix des constantes l_1 et l_2. Il n'a pas la prétention d'être une preuve rigoureuse.

Si on se place dans le cadre de l'hypothèse "A_+", d'après les équations des filtres de Kalman-Bucy (2.3) et (2.4), on a:

$$\begin{aligned} Z_{k+1} &= (1+\varepsilon B_-)(1-H_-K_-)\,Z_k + \varepsilon\,(1+\varepsilon B_+)(B_+ - B_-)H_+\hat{x}_k^+ \\ &\quad +(H_+K_+(1+\varepsilon B_+) - H_-K_-(1+\varepsilon B_-))\nu_{k+1}^+. \end{aligned} \tag{4.3}$$

Remarque 4.2 Sous (HDi), $H_+K_+ - H_-K_-$ est d'ordre $\mathcal{O}(1)$.

Vu que les filtres en question sont à "mémoire courte", d'après l'expression (4.3), l'hypothèse d'indépendance sur A^+ de ν_k^+ et ν_{k+1}^+, nous permet de supposer "l'indépendance" de Z_k et ν_{k+1}^+ à partir d'un certain instant $i_1 > i_0$.

Remarque 4.3 On a la majoration

$$(1+\varepsilon B_-)(1-H_-K_-) < 1\,.$$

Pour un intervalle de monotonie suffisament grand, nous supposons que la situation stationnaire est atteinte. Dans le cas où $B_+ \neq B_-$ nous émettons donc l'hypothèse supplémentaire:

$(H4)$: $\quad B_+ < 0,\, B_- < 0.$

On obtient alors

- si $B_+ = B_- \triangleq B$, dans une situation asymptotique,

$$\mathbf{E}[Z^2\,;A_+] = \varepsilon\,\frac{(1+\varepsilon B)^2(H_+K_+ - H_-K_-)^2}{1-(1+\varepsilon B)^2(H_+K_+ - H_-K_-)^2}\,\vartheta_+\,.$$

- si $B_+ < 0$, $\mathbf{E}[\hat{x}_k^{+2}\,;A_+]$ admet une solution d'équilibre:

$$\mathbf{E}[\hat{x}_k^{+2}\,;A_+] = \frac{K_+^2}{|B_+|(2+\varepsilon B_+)}\,\vartheta^+\,.$$

En utilisant l'inégalité

$$\mathbf{E}[Z_k\,\hat{x}_k^+\,;A_+] \le \frac{1}{2\sqrt{\epsilon}}\,\mathbf{E}[Z_k^2\,;A_+] + \frac{\sqrt{\epsilon}}{2}\,\mathbf{E}[\hat{x}_k^+\,;A_+])\,,$$

et la majoration suivante, pour ϵ "assez petit",

$$\begin{aligned}(1+\epsilon B_-)^2(1-H_-K_-)^2 + \sqrt{\epsilon}\,H_+|B_+-B_-|(1-H_-K_-)(1+\epsilon B_+)(1+\epsilon B_-)\\ < 1 - H_-K_-(2-H_-K_-) + c\sqrt{\epsilon} < 1\,,\end{aligned}$$

on obtient l'estimation

$$\begin{aligned}&\mathbf{E}[Z_{k+1}^2\,;A_+]\\ \le\ &(1-H_-K_-(2-H_-K_-)+c\sqrt{\epsilon})\mathbf{E}[Z_k^2\,;A_+]\\ &+\epsilon^{3/2}(1+\epsilon B_+)(1+\epsilon B_-)|B_+-B_-|(1-H_-K_-)H_+\mathbf{E}[\hat{x}^{+2}\,;A_+]\\ &+\epsilon\,[H_+K_+(1+\epsilon B_+)-H_-K_-(1+\epsilon B_-)]^2\vartheta^+\\ &+\epsilon^2(1+\epsilon B_+)^2(B_+-B_-)^2H_+^2\mathbf{E}[\hat{x}^{+2}\,;A_+]\,.\end{aligned}$$

Dans une situation stationnaire, on a

$$\mathbf{E}[Z^2\,;A_+] \le \Lambda_+\,\epsilon + c\epsilon^{3/2}\,,$$

où

$$\Lambda_+ = \frac{(H_+K_+ - H_-K_-)^2}{H_-K_-(2-H_-K_-)}\,\vartheta_+\,.$$

De plus, pour un certain n_0, on peut considérer Z_k comme une combinaison linéaire de $\nu_{k-n_0}^+,\cdots,\nu_k^+$. Le processus $\xi_k = Z_k\nu_{k+1}^+$ est alors approximativement un processus mélangeant. En conséquence, on approche $\epsilon\hat{L}_n$ par le processus de diffusion $\{\zeta_t;\ t=\epsilon n\}$ solution de l'équation différentielle stochastique

$$d\zeta_t = \Lambda_+/2\,dt + \sqrt{\epsilon\Lambda_+\vartheta^+}\,dW_t^+.$$

Comme pour le test de la variation quadratique, on obtient la probabilité d'erreur $p_+ = P(\{\zeta_{T_+^*} = -\epsilon l_1\}|A_+)$:

$$p_+ = \frac{1-e^{-\eta_+ l_2}}{e^{\eta_+ l_1} - e^{-\eta_+ l_2}}\,,$$

où T_+^* est le temps d'arrêt défini par

$$T_+^* = \inf\{t\,:\,\zeta_t \ge \epsilon\,l_2 \text{ ou } \zeta_t \le -\epsilon\,l_1\}$$

et $\eta_+ = 1/\vartheta^+$.

Si on se place dans le cadre de l'hypothèse "A_-", par un raisonnement identique au précédent, on obtient la probabilité d'erreur $p_- = P(\{\zeta_{T_-^*} = \varepsilon l_2\}|A_-)$. Elle est donnée par:

$$p_- = \frac{1 - e^{-\eta_- l_1}}{e^{\eta_- l_2} - e^{-\eta_- l_1}},$$

avec $\eta_- = 1/\vartheta^-$.

Pour des valeurs de p_+ et p_- données ([4]), on détermine les bornes de précision l_1 et l_2 et on estime les temps moyens pour prendre une décision:

$$\begin{aligned} \mathbf{E}(T_+^*) &= 2\varepsilon\left(\frac{l_2}{\Lambda_+} - p_+\frac{l_1 + l_2}{\Lambda_+}\right), \\ \mathbf{E}(T_-^*) &= 2\varepsilon\left(\frac{l_1}{\Lambda_-} - p_-\frac{l_1 + l_2}{\Lambda_-}\right). \end{aligned}$$

Remarque 4.4 Dans [4], les auteurs constatent qu'en pratique le temps moyen pour prendre une décision est plus long pour le test du rapport de vraisemblance que pour le test de la variation quadratique.

[4] En général on prend $p_- = p_+ = 0.05$.

5 Sous l'hypothèse $(HD2)$: décider du signe de x_k

Sous l'hypothèse $(HD2)$: $\Gamma^2 = (H_-\sigma_-)^2 = (H_+\sigma_+)^2$ et $B_- \neq B_+$, le test de la variation quadratique n'est pas applicable. En remplacement nous considérons un test du type rapport de vraisemblance sur les observations y_k. Puis nous adapterons, sous cette hypothèse, le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy, présenté dans le paragraphe précédent sous l'hypothèse $(HD1)$.

5.1 Test du rapport de vraisemblance sur les acroissements des observations

Ce test est déduit de l'étude du cas $\varepsilon = 0$.

Le cas $\varepsilon = 0$

On considère le système suivant:

$$\begin{cases} x_{k+1} &= x_k + \Delta t\, b(x_k) + \sqrt{\Delta t}\, \sigma(x_k)\, u_k, \\ \tilde{y}_k &= h(x_k)\,. \end{cases} \tag{5.1}$$

L'équation d'état de ce système généralise celle du système discret (1.4) à un pas de temps Δt .

Le problème est de déterminer le signe de x_k pour $k = i_0, i_0+1, \cdots, i_0+m$. Le test d'hypothèses est fait sur les observations $\{\tilde{y}_k, i_0+1 \leq k \leq i_0+m\}$.

On considère la filtration $(\mathcal{F}_{i_0}^k)_k$, où $\mathcal{F}_{i_0}^k = \sigma\{x_{i_0}, u_{i_0}, \cdots, u_k\}$, $i_0 \leq k \leq i_0+m$.

$$\text{Sur } A_+ : \qquad \hat{y}_{k+1} = (1 + \Delta t B_+)\tilde{y}_k + \sqrt{\Delta t}\,\Gamma u_k^+ . \tag{5.2}$$

où $\{u_k^+;\, k \geq i_0\}$ est un $\mathcal{F}_{i_0}^k$-bruit blanc gaussien pour une probabilité P^+ telle que $P^+(D) = P(D), \forall D \subset A_+$.

$$\text{Sur } A_- : \qquad \hat{y}_{k+1} = (1 + \Delta t B_-)\tilde{y}_k + \sqrt{\Delta t}\,\Gamma u_k^- . \tag{5.3}$$

où $\{u_k^-;\, k \geq i_0\}$ est un $\mathcal{F}_{i_0}^k$-bruit blanc gaussien pour une probabilité P^- telle que $P^-(D) = P(D), \forall D \subset A_-$.

Le processus $\{\tilde{y}_k;\, k \geq i_0+1\}$ est $\mathcal{F}_{i_0}^{k-1}$-mesurable (i.e. prévisible) et $\sum_{k=i_0}^{i_0+m} \tilde{y}_k^2 < +\infty$ P-p.s.

On utilise une version discrète du théorème de Girsanov (cf. [10]) et, vu que $B_- \neq B_+$, on prend pour statistique le logarithme du rapport de vraisemblance:

$$L_n = \ln \frac{dP^+}{dP^-}\Big|_{\mathcal{F}_{i_0}^{i_0+n}},$$

où $1 \leq n \leq m-1$. On a la formule:

$$L_n = \frac{1}{\Gamma^2}\left[(B_+ - B_-)\sum_{k=i_0+1}^{i_0+n} \tilde{y}_k(\tilde{y}_{k+1} - \tilde{y}_k) - \frac{1}{2}\Delta t\,(B_+^2 - B_-^2)\sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2\right]. \qquad (5.4)$$

Sur l'événement A_+, la statistique L_n se reécrit sous la forme:

$$L_n = +\frac{1}{2}\Delta t\,\frac{(B_+ - B_-)^2}{\Gamma^2}\sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 + \sqrt{\Delta t}\,\frac{B_+ - B_-}{\Gamma}\sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^+. \qquad (5.5)$$

Sur l'évènement A_-, on a:

$$L_n = -\frac{1}{2}\Delta t\,\frac{(B_+ - B_-)^2}{\Gamma^2}\sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 + \sqrt{\Delta t}\,\frac{B_+ - B_-}{\Gamma}\sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^-. \qquad (5.6)$$

On pose

$$R_n = \Delta t\,\frac{(B_+ - B_-)^2}{\Gamma^2}\sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2.$$

On appelle R_n le terme de précision. Pour R_n grand, L_n/R_n est proche de $\pm 1/2$ et la probabilité d'erreur est faible. Si $L_n \geq 0$ on décide A_+ et si $L_n < 0$ on décide A_-.

On établit le résultat suivant:

Théorème 5.1 *Soit $r > 0$. $\forall n$ tel que $i_0 < n \leq m-1$, on a:*

$$\begin{aligned} P(\{L_n < 0\} \cap \{R_n \geq r\} \cap A_+\}) &\leq e^{-r/8}, \\ P(\{L_n \geq 0\} \cap \{R_n \geq r\} \cap A_-\}) &\leq e^{-r/8}. \end{aligned}$$

Preuve La démonstration est identique pour les deux inégalités. Considérons par exemple la première.

On introduit la martingale exponentielle

$$Z_k = \exp\left\{\lambda\sqrt{\Delta t}\,\frac{B_+ - B_-}{\Gamma}\sum_{i=i_0+1}^{i_0+k} \tilde{y}_i\,u_i^+ - \frac{1}{2}\lambda^2\Delta t\,\frac{(B_+ - B_-)^2}{\Gamma^2}\sum_{i=i_0+1}^{i_0+k} \tilde{y}_i^2\right\},$$

avec $\lambda < 0$ et on raisonne comme dans le cas continu (cf. chapitre I [prop.5.5]). □

Les calculs qui suivent utilisent le fait qu'une situation stationnaire est nécéssairement atteinte par $\{\tilde{y}_k\}$, raison pour laquelle nous émettons l'hypothèse supplémentaire:

(H4): $B_+ < 0$ et $B_- < 0$.

On décrit deux procédures de mise en oeuvre de ce test sous forme séquentielle.

Première procédure:

Pour r fixé, on introduit le temps d'arrêt N^*:

$$N^* = \inf\{n \geq 1 \,:\, R_n \geq r\} \wedge m\,.$$

- Si $N^* \leq m$:
  - si $L_{N^*} \geq 0$ on décide A_+,
  - si $L_{N^*} < 0$ on décide A_-;
- Sinon on ne peut pas décider.

Nous devons donc proposer une méthode de détermination de la constante r. On considère le problème sur A_+ et sur A_- et on choisit $r = \max(r_-, r_+)$.

Sur l'évènement A_+, on pose

$$M_n^+ = \sqrt{\Delta t}\,\frac{B_+ - B_-}{\Gamma} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k\, u_k^+\,.$$

La statistique L_n se réécrit sous la forme:

$$L_n = \frac{1}{2} R_n + M_n^+\,.$$

On introduit le temps d'arrêt N_+^*:

$$N_+^* = \inf\{n \geq 1 \,:\, R_n \geq r_+\}\,.$$

On a que $R_{N_+^*}$ est approximativement égal à r_+.

On cherche à déterminer une constante r_+ telle que, pour une probabilité $1-\alpha$ donnée, le terme de précision soit significatif dans le calcul de $L_{N_+^*}$. Dans d'autres termes, on cherche r_+ telle que:

$$P(\{-M_{N_+^*}^+ > \frac{1}{2} r_+\} | A_+) \leq \alpha\,. \tag{5.7}$$

D'après (5.2), on réécrit $\tilde{y}_k$ sous la forme suivante:

$$\tilde{y}_{k+1} = (1 + \Delta t B_+)^{n_0} \tilde{y}_{k+1-n_0} + \sqrt{\Delta t}\,\Gamma \sum_{i=0}^{n_0-1} (1 + \Delta t B_+)^i u_{k-i}^+\,.$$

L'hypothèse $(H4)$ nous permet de considérer $\tilde{y}_k$ comme une combinaison linéaire des variables aléatoires gaussiennes et indépendantes $u^+_{k-n_0+1}, \cdots, u^+_k$. Soit $\xi_k = \tilde{y}_k u^+_k$. Le processus $\{\xi_k, i_0+n_0-1 \leq k \leq i_0+n\}$ est alors approximativement un processus mélangeant. De plus, la valeur stationnaire de $E(\tilde{y}^2_k\,;\, A_+)$ est approchée par

$$\Lambda_+ = \frac{\Gamma^2}{-B_+(2+\Delta t B_+)}.$$

D'après le théorème fonctionnel de limite centrale [2][th.20.1],

$$\frac{\sqrt{\Delta t}}{\sqrt{\Lambda_+}} \sum_{k=i_0+1}^{i_0+n} \xi_k \approx W_t,$$

où $t = n\Delta t$ et $\{W_t\}$ est un processus de Wiener standard.
On déduit de ce qui précède que M^+_n suit approximativement une loi normale

$$\mathcal{N}\left(0, n\Delta t \frac{(B_+ - B_-)^2}{-B_+(2+\Delta t B_+)}\right).$$

Vu que $R_{N^*_+} \approx r_+$, on obtient la relation suivante entre r_+ et N^*_+ :

$$r_+ \approx N^*_+ \Delta t \, \frac{(B_+ - B_-)^2}{-B_+(2+\Delta t B_+)}.$$

Alors l'inégalité (5.7) devient:

$$P\left(\{\frac{-M^+_{N^*_+}}{\sqrt{r_+}} > \frac{1}{2}\sqrt{r_+}\}|A_+\right) \leq \alpha.$$

On choisit

$$r_+ = 4\,\lambda^2,$$

avec λ obtenu par la table de la loi normale centrée réduite.

Sur l'évènement A_-, on prend

$$r_- = 4\,\lambda^2$$

et on a la relation suivante entre r_- et N^*_- :

$$r_- \approx N^*_- \Delta t \, \frac{(B_+ - B_-)^2}{-B_-(2+\Delta t B_-)}.$$

Exemple 5.2 Pour $B_- = -1$, $B_+ = -0.25$ et $\alpha = 0.05$, on obtient $r = 10.89$. Pour les estimations des temps moyens, on obtient: $\Delta t N^*_+ = 9.66$, $\Delta t N^*_- = 38.52$.

Deuxième procédure:

Il semble que dans le test précédent, le temps d'attente pour prendre une décision soit "trop" long. En effet les bornes r_+ et r_- sont déterminées indépendamment l'une de l'autre. Pour réduire cette durée, nous proposons une autre procédure.

Pour des constantes $l_1 > 0$, $l_2 > 0$ à déterminer, on introduit le temps d'arrêt N^*:

$$N^* = \inf\{n > i_0 \text{ t.q. } L_n \geq l_2 \text{ ou } L_n \leq -l_1\} \wedge m\,.$$

Règle de décision :

- Si $L_{N^*} \geq l_2$ on décide "A_+".
- Si $L_{N^*} \leq -l_1$ on décide "A_-".
- Sinon, on ne peut pas décider.

On utilise à nouveau l'approximation des processus discrets par des diffusions pour déterminer les constantes l_1 et l_2.

Sur A_+, on approche L_n par $\{\zeta_t;\ t = \Delta t\, n\}$, où $\{\zeta_t\}$ est le processus de diffusion défini par l'équation différentielle stochastique

$$d\zeta_t = \frac{(B_+ - B_-)^2 \Lambda_+}{2\,\Gamma^2}\,dt + \frac{\sqrt{\Lambda_+}\,(B_+ - B_-)}{\Gamma}\,dW_t\,.$$

On introduit le temps d'arrêt T_+^*:

$$T_+^* = \inf\{t \geq 0 :\ \zeta_t \geq l_2 \text{ ou } \zeta_t \leq -l_1\}\,.$$

On définit la probabilité $p_+(x) = P_{\zeta_0 = x}(\{\zeta_{T_+^*} = -l_1\}|A_+)$. En utilisant la formule de Dynkin, on obtient l'équation différentielle ordinaire suivante:

$$\begin{cases} p_+'' + p_+' = 0 \\ p_+(-l_1) = 1\,,\ p_+(l_2) = 0\,. \end{cases}$$

La résolution de cette équation nous donne la probabilité de "refuser A_+ à tort ", $p_+ = p_+(0)$ (cf. [6][chapV.th7.1]):

$$p_+ = \frac{1 - e^{-l_2}}{e^{l_1} - e^{-l_2}}\,.$$

Par un raisonnement identique sur A_-, on obtient l'expression de la probabilité d'erreur $p_- = p_-(0)$:

$$p_- = \frac{1 - e^{-l_1}}{e^{l_2} - e^{-l_1}}\,.$$

De même, on calcule les temps moyens pour prendre une décision en résolvant deux équations différentielles ordinaires. (cf. [6][chapV.th7.1]). Plus précisément $\mathbf{E}(T_+^*)$ est donné par la solution de l'équation

$$\begin{cases} \dfrac{\Lambda_+ (B_+ - B_-)^2}{2\,\Gamma^2}\, g_+'' + \dfrac{\Lambda_+ (B_+ - B_-)^2}{2\,\Gamma^2}\, g_+' + 1 = 0 \\ g_+(-l_1) = g_+(l_2) = 0 \end{cases}$$

au point 0.

On obtient $\mathbf{E}(T_-^*)$ de façon similaire:

$$\begin{aligned} \mathbf{E}(T_+^*) &= \frac{-2\,B_+(2 + \Delta t\, B_+)}{(B_+ - B_-)^2}\,[l_2 - p_+(l_1 + l_2)]\,, \\ \mathbf{E}(T_-^*) &= \frac{-2\,B_-(2 + \Delta t\, B_-)}{(B_+ - B_-)^2}\,[l_1 - p_-(l_1 + l_2)]\,. \end{aligned}$$

Remarque 5.3 Dans les applications, on prend généralement $p_+ = p_- = \alpha$. Ainsi un simple calcul nous donne

$$l_1 = l_2 = \ln \frac{1-\alpha}{\alpha}\,.$$

et on note que

$$\mathbf{E}(T_-^*) \approx \frac{B_-}{B_+}\,\mathbf{E}(T_+^*)\,.$$

Exemple 5.4 Pour $B_- = -1$, $B_+ = -0.25$ et $\alpha = 0.05$, on obtient $l_1 = l_2 = 2.94$. Les temps moyens sont alors $E(T_+^*) = 4.70$ et $E(T_-^*) = 18.79$.

Remarque 5.5 Pour comparer les temps moyens des deux procédures, on s'intéresse au signe de l'expression:

$$2\,\lambda^2 - (1 - 2\,\alpha)\ln \frac{1-\alpha}{\alpha}\,,$$

avec λ tel que $\dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\lambda} e^{-x^2/2} dx = \alpha$.

On vérifie numériquement que cette expression est positive, $\forall\, 0 < \alpha < 1/2$. On en déduit que pour une probabilité d'erreur donnée, les temps moyens de la première procédure sont plus grands que ceux de la deuxième. Cette différence est d'autant plus grande que α est petit.

En préparation de l'étude du cas $\epsilon > 0$, nous réécrivons la statistique L_n sous une autre forme. En utilisant l'égalité

$$\hat{y}_k\,(\hat{y}_{k+1} - \hat{y}_k) = -\frac{1}{2}\left((\hat{y}_{k+1} - \hat{y}_k)^2 - \hat{y}_{k+1}^2 + \hat{y}_k^2\right)$$

on obtient

$$L_n = \frac{B_+ - B_-}{2\Gamma^2}(\tilde{y}^2_{i_0+n+1} - \tilde{y}^2_{i_0+1}) - \frac{B_+ - B_-}{2\Gamma^2}\sum_{k=i_0+1}^{i_0+n}(\tilde{y}_{k+1} - \tilde{y}_k)^2$$
$$-\Delta t\,\frac{B_+^2 - B_-^2}{2\Gamma^2}\sum_{k=i_0+1}^{i_0+n}\tilde{y}_k^2 .$$

Lemme 5.6 *Sur $A_+ \cup A_-$, on a l'estimation:*

$$\sum_{k=i_0+1}^{i_0+n}(\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n\,\Delta t = \mathcal{O}(\sqrt{\Delta t}) .$$

Preuve Démontrons, par exemple, ce résultat sur l'évènement A_+.

Sur A_+, on a :

$$\tilde{y}_{k+1} = \tilde{y}_k + \Delta t\, H_+ B_+ x_k + \sqrt{\Delta t}\,\sigma_+ H_+ u_k ,$$

d'où l'expression:

$$\sum_{k=i_0+1}^{i_0+n}(\tilde{y}_{k+1} - \tilde{y}_k)^2 = \Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} x_k^2 + \Delta t\,\Gamma^2 \sum_{k=i_0+1}^{i_0+n} u_k^2$$
$$+2\,\Delta t^{3/2} H_+^2 \sigma_+ B_+ \sum_{k=i_0+1}^{i_0+n} x_k u_k .$$

On obtient l'estimation:

$$\mathbf{E}\left[\sum_{k=i_0+1}^{i_0+n}(\tilde{y}_{k+1} - \tilde{y}_k)^2 ;\, A_+\right] \le \Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n}\mathbf{E}(x_k^2) + \Delta t\, n\,\Gamma^2 .$$

L'égalité

$$\mathbf{E}\left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2\right)\left(\sum_{k=i_0+1}^{i_0+n} x_k u_k\right)\right] = 0 ,$$

entraîne

$$\mathbf{E}\left[\left(\sum_{k=i_0+1}^{i_0+n}(\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n\,\Delta t\right)^2 ;\, A_+\right]$$
$$\le \mathbf{E}\left[\left(\Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} x_k^2 + 2\,\Delta t^{3/2} H_+^2 \sigma_+ B_+ \sum_{k=i_0+1}^{i_0+n} x_k u_k\right)^2\right]$$

$$+\Delta t^2\Gamma^4\mathbf{E}\left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2\right)^2\right] + 2\,\Delta t^3\Gamma^2\,H_+^2\,B_+^2\mathbf{E}\left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2\right)\left(\sum_{k=i_0+1}^{i_0+n} x_k^2\right)\right]$$

$$-\Delta t^2\,n^2\,\Gamma^4 - 2\,\Delta t^3\,n\,\Gamma^2\,H_+^2\,B_+^2\sum_{k=i_0+1}^{i_0+n}\mathbf{E}(x_k^2)\,. \tag{5.8}$$

On remarque que

$$\mathbf{E}\left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2\right)^2\right] = \sum_{k=i_0+1}^{i_0+n}\mathbf{E}(u_k^4) + n(n-1) = n(n+2)\,,$$

puisque $\{u_k\}$ est un bruit blanc standard. Ainsi on a:

$$\Delta t^2\Gamma^4\mathbf{E}\left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2\right)^2\right] - \Delta t^2\,n^2\,\Gamma^4 = 2\Delta t^2\,n\,\Gamma^4$$

On utilise l'inégalité $(a+b)^2 \le 2a^2 + 2b^2$ pour traiter le premier terme de (5.8) et l'inégalité de Cauchy - Schwartz pour le troisième. Vu que les moments d'ordre 2 et 4 de x_k sont finis sur l'intervalle $[a,b]$ et que $n \le (b-a)/\Delta t$, on obtient la majoration suivante:

$$\mathbf{E}\left[\left(\sum_{k=i_0+1}^{i_0+n}(\tilde{y}_{k+1}-\tilde{y}_k)^2 - \Gamma^2\,n\,\Delta t\right)^2\;;\,A_+\right] \le c_{a,b}\,\Delta t\,,$$

où $c_{a,b}$ est une constante positive dépendante de l'intervalle $[a,b]$ fixé.

Sur l'événement A_-, la démonstration est similaire.

□

Le cas $\varepsilon > 0$

On définit la statistique L_n^ε comme il suit:

$$\begin{aligned} L_n^\varepsilon &= \frac{B_+ - B_-}{2\,\Gamma^2}\,(y_{i_0-n+1}^2 - y_{i_0+1}^2) - \frac{1}{2}\,n\,\varepsilon\,(B_+ - B_-) \\ &\quad -\varepsilon\,\frac{B_+^2 - B_-^2}{2\,\Gamma^2}\sum_{k=i_0+1}^{i_0+n} y_k^2\,. \end{aligned} \tag{5.9}$$

On considère $\Delta t = \varepsilon$ et on s'intéresse à l'écart $L_n^\varepsilon - L_n$.

Théorème 5.7 *Sur les événements C et $\hat{C}$ définis dans le paragraphe 3, on a:*

$$L_n^\varepsilon - L_n = \mathcal{O}(\sqrt{\varepsilon})\,.$$

Preuve

On démontre d'abord que $L_n^\epsilon - L_n$ est d'ordre $\mathcal{O}(\sqrt{\epsilon})$ sur $A_+ \cup A_-$.

$$\begin{aligned} L_n^\epsilon - L_n &= \frac{B_+ - B_-}{2\,\Gamma^2}\left((y_{i_0+n}^2 - \tilde{y}_{i_0+n}^2) - (y_{i_0+1}^2 - \tilde{y}_{i_0+1}^2)\right) \\ &\quad + \frac{B_+ - B_-}{2\,\Gamma^2}\left[\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2\, n\,\epsilon\right] \\ &\quad - \epsilon\,\frac{B_+^2 - B_-^2}{2\,\Gamma^2}\sum_{k=i_0+1}^{i_0+n}(y_k^2 - \tilde{y}_k^2)\,. \end{aligned}$$

L'égalité

$$y_k^2 - \tilde{y}_k^2 = \sqrt{\epsilon}\,v_k\,\left(2\,h(x_k) + \sqrt{\epsilon}\,v_k\right)$$

implique

$$\begin{aligned} L_n^\epsilon - L_n &= \sqrt{\epsilon}\,\frac{B_+ - B_-}{\Gamma^2}\,(v_{i_0+n+1}\,h(x_{i_0+n+1}) - v_{i_0+1}\,h(x_{i_0+1})) \\ &\quad + \frac{B_+ - B_-}{2\,\Gamma^2}\left[\sum_{k=i_0+1}^{i_0+n}(h(x_{k+1}) - h(x_k))^2 - \Gamma^2\,n\,\epsilon\right] \\ &\quad + \varepsilon\,\frac{B_+ - B_-}{2\,\Gamma^2}(v_{i_0+n+1}^2 + v_{i_0+1}^2) - \epsilon^{3/2}\,\frac{B_+^2 - B_-^2}{2\,\Gamma^2}\sum_{k=i_0+1}^{i_0+n} v_k\,h(x_k) \\ &\quad + \epsilon^2\,\frac{B_+^2 - B_-^2}{2\,\Gamma^2}\sum_{k=i_0+1}^{i_0+n} v_k^2\,. \end{aligned}$$

Sachant que $\{v_k\}$ est un bruit blanc gaussien standard et que tous les moments de x_k sont finis pour $k \le m$, on déduit à l'aide du lemme 5.6 que $L_n^\epsilon - L_n$ est d'ordre $\mathcal{O}(\sqrt{\epsilon})$ sur $A_+ \cup A_-$.

Pour raisonner sur $\mathcal{C}$, on utilise la majoration

$$\begin{aligned} \mathbf{E}[(L_n^\epsilon - L_n)^2\,;\,\mathcal{C}] &\le \mathbf{E}[(L_n^\epsilon - L_n)^2\,;\,A_+ \cup A_-] \\ &\quad + \mathbf{E}[(L_n^\epsilon - L_n)^2\,;\,\mathcal{C} \cap (A_+ \cup A_-)^c]\,. \end{aligned}$$

De l'inégalité de Cauchy-Schwartz, on obtient

$$\mathbf{E}[(L_n^\epsilon - L_n)^2\,;\,\mathcal{C}] \le \mathbf{E}[(L_n^\epsilon - L_n)^4]^{1/2}\,P[\mathcal{C} \cap (A_+ \cup A_-)^c]^{1/2} + c\,\epsilon\,.$$

A l'aide du lemme 3.4, on montre que $\mathbf{E}[(L_n^\epsilon - L_n)^4] \le c$ et du théorème 3.1, on déduit le résultat sur $\mathcal{C}$.

De même, du théorème 3.7, on déduit le résultat sur $\hat{\mathcal{C}}$. □

Etant donné que les bornes déterminées dans le paragraphe 5.1 sont d'ordre $\mathcal{O}(1)$, il est donc légitime d'après le résultat précédent de les utiliser pour tester L_n^ε.

D'aprés la remarque 5.5, nous avons choisi de mettre en œuvre le test selon sa deuxième version.

5.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman–Bucy

On montre que la statistique $\hat{L}_n$, utilisée dans le paragraphe 4.2 sous l'hypothèse $(HD1)$, permet encore de décider entre les alternatives "A_+" et "A_-" sous l'hypothèse $(HD2)$.

Soit $i_0 < i_1 < i_0 + m$. On a

$$\hat{L}_n = \frac{1}{\varepsilon}\sum_{k=i_1}^{i_1+n} Z_k\, y_{k+1} - \frac{1}{2\varepsilon}\sum_{k=i_1}^{i_1+n}\left((H_+(1+\varepsilon B_+)\hat{x}_k^+)^2 - (H_-(1+\varepsilon B_-)\hat{x}_k^-)^2\right),$$

où $Z_k = H_+(1+\varepsilon B_+)\hat{x}_k^+ - H_-(1+\varepsilon B_-)\hat{x}_k^-$.

On rappelle la définition du terme de précision du test:

$$\hat{R}_n = \frac{1}{\varepsilon}\sum_{k=i_1}^{i_1+n} Z_k^2.$$

Sur A_+, à partir de l'expression récurrente (4.3) de Z_k, on obtient

$$\begin{aligned}
&\mathbf{E}[Z_{k+1}^2\,;A_+]\\
\leq\ &(1+\varepsilon B_-)^2(1-H_-K_-)^2\mathbf{E}[Z_k^2\,;A_+]\\
&+\varepsilon^2(1+\varepsilon B_+)^2(B_+-B_-)^2H_+^2\mathbf{E}[\hat{x}^{+2}\,;A_+]\\
&+\varepsilon\,[H_+K_+(1+\varepsilon B_+)-H_-K_-(1+\varepsilon B_-)]^2\vartheta^+\\
&+2\varepsilon\,(1+\varepsilon B_+)(1+\varepsilon B_-)|B_+-B_-|(1-H_-K_-)H_+\mathbf{E}[Z_k\,\hat{x}_k^+\,;A_+]
\end{aligned}$$

Remarque 5.8 Sous $(HD2)$, la différence $H_+K_+ - H_-K_-$ est d'ordre $\mathcal{O}(\varepsilon)$, alors que sous $(HD1)$, elle était d'ordre $\mathcal{O}(1)$.

On utilise l'inégalité

$$\begin{aligned}
\mathbf{E}[Z_k\,\hat{x}_k^+\,;A_+] \leq\ &\frac{H_-K_-}{2\varepsilon(1+\varepsilon B_-)(1+\varepsilon B_+)|B_+-B_-|H_+}\mathbf{E}[Z_k^2\,;A_+]\\
&+\frac{\varepsilon(1+\varepsilon B_-)(1+\varepsilon B_+)|B_+-B_-|H_+}{2H_-K_-}\mathbf{E}[\hat{x}^{+2}\,;A_+],
\end{aligned}$$

et les majorations suivantes, pour ε suffisament petit,

$$\begin{aligned} 0 &< (1+\varepsilon B_+)^2(1-H_-K_-)^2 + H_-K_-(1-H_-K_-) \\ &< 1 - H_-K_- + c\varepsilon < 1\,. \end{aligned}$$

De la même façon que dans le paragraphe 4.2, sous l'hypothèse $(H4)$, on obtient

$$\mathbf{E}[Z^2\,; A_+] \leq \hat{\Lambda}_+\,\varepsilon^2 + c\varepsilon^3\,,$$

où

$$\hat{\Lambda}_+ = \frac{(B_+ - B_-)^2}{|B_+|(2+\varepsilon B_+)}\,\vartheta_+\,.$$

Sur A_-, on fait un raisonnement similaire.

Remarque 5.9 Le processus $\{Z_k\}$ est d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ sous l'hypothèse $(HD1)$ et d'ordre $\mathcal{O}(\varepsilon)$ sous l'hypothèse $(HD2)$.

De même que pour le test précédent, on considère deux types de procédures.

Première procédure:

Pour r fixé, on calcule $\hat{L}_n$ jusqu'à ce qu'on ait $\hat{R}_n \geq r$. Par un raisonnement similaire à celui du paragraphe 5.1, on choisit $r = \max(r_-, r_+)$ avec

$$r_- = 4\lambda^2\vartheta^- \text{ et } r_+ = 4\lambda^2\vartheta^+.$$

La constante λ est obtenue par inversion de la fonction de répartition de la loi $\mathcal{N}(0,1)$. Les estimations des temps moyens pour prendre une décision sont données par

$$\begin{aligned} \varepsilon N_-^* &= 4\lambda^2 \frac{-B_-(2+\varepsilon B_-)}{(B_+ - B_-)^2}, \\ \varepsilon N_+^* &= 4\lambda^2 \frac{-B_+(2+\varepsilon B_+)}{(B_+ - B_-)^2}. \end{aligned}$$

Remarque 5.10 Les temps moyens pour prendre une décision associés à cette procédure sont identiques à ceux du test précédent quand on considère le même type d'approche (cf. paragraphe 5.1).

Deuxième procédure:

Soit $l_1 > 0$ et $l_2 > 0$ des bornes fixées, on introduit le temps d'arrêt N^*:

$$N^* = \inf\{n > i_1 \,:\, \hat{L}_n \geq l_2 \text{ ou } \hat{L}_n \leq -l_1\} \wedge m\,.$$

Raisonnons sur l'événement A_+, par exemple.
Nous savons, d'après ce qui précède, que sous l'hypothèse $(H4)$, le processus $\xi_k = Z_k \nu_{k+1}$ est approximativement un processus mélangeant. Par un raisonnement similaire à celui du paragraphe 4.2, pour $t = \varepsilon n$, on approche $\hat{L}_n$ par le processus de diffusion $\{\zeta_t\}$ défini par:

$$d\zeta_t = \hat{\Lambda}_+/2\, dt + \sqrt{\vartheta^+ \hat{\Lambda}_+}\, dW_t^+\,.$$

On obtient la probabilité d'erreur p_+, "probabilité de refuser A_+ à tort":

$$p_+ = \frac{1 - e^{-\eta_+ l_2}}{e^{\eta_+ l_1} - e^{-\eta_+ l_2}},$$

où $\eta_+ = 1/\vartheta^+$ ainsi que le temps moyen pour prendre une décision sur A_+ :

$$E(T_+^*) = 2\left(\frac{l_2}{\hat{\Lambda}_+} - p_+ \frac{(l_1 + l_2)}{\hat{\Lambda}_+}\right).$$

De même, on obtient la probabilité d'erreur p_-, "probabilité de refuser A_- à tort":

$$p_- = \frac{1 - e^{-\eta_- l_1}}{e^{\eta_- l_2} - e^{-\eta_- l_1}},$$

où $\eta_- = 1/\vartheta^-$ et le temps moyen pour prendre une décision sur A_- :

$$E(T_-^*) = 2\left(\frac{l_1}{\hat{\Lambda}_-} - p_- \frac{(l_1 + l_2)}{\hat{\Lambda}_-}\right).$$

Remarque 5.11 Sous $(HD1)$, les temps d'attente moyens sont d'ordre $\mathcal{O}(\varepsilon)$ tandis que sous $(HD2)$, ils sont d'ordre $\mathcal{O}(1)$. Dans ce deuxième cas, on peut donc s'attendre à ce qu'il y ait des intervalles où aucune décision n'est prise.

Remarque 5.12 Sous l'hypothèse $(HD2)$, l'écart $\vartheta^+ - \vartheta^-$ est d'ordre $\mathcal{O}(\varepsilon)$, donc les bornes l_1 et l_2 sont "très proches". On a une situation similaire à celle décrite dans la remarque 5.5, i.e. les temps moyens d'attente sont plus longs dans la première procédure que dans la deuxième.

Le test du rapport de vraisemblance sur les sorties des filtres de Kalman sera mis en œuvre selon la deuxième procédure.

6 Application

Nous rappelons que notre objectif est d'étudier le comportement des filtres proposés par rapport à la solution optimale du problème de filtrage et de comparer les différentes procédures permettant de les obtenir. Nous devons donc, en premier lieu, définir des critères de comparaison.

6.1 Les critères de comparaison

Une première idée est de considérer, pour une probabilité d'erreur fixée,

- l'amplitude moyenne des intervalles détectés, pour comparer les tests de détection
- le temps moyen nécessaire pour prendre une décision, pour comparer les tests de détermination du signe.

Pour comparer les tests en termes de probabilité d'erreur, il semble naturel de s'intéresser aux pourcentages suivants:

- pourcentage d'intervalles détectés corrects.
- pourcentage de décisions correctes (i.e. nombre de décisions correctes par rapport au nombre de décisions prises).

Nous qualifions une décision de correcte lorsque le signe choisi correspond à celui de la trajectoire sur un intervalle de monotonie correct.

Les temps moyens d'attente pour prendre une décision sont calculés d'abord théoriquement (voir les paragraphes 4 et 5) et sont ensuite confrontés aux résultats obtenus par simulation.

6.2 Exemples

L'étude du problème sous l'hypothèse $(HD2)$ étant notre principal centre d'intérêt, nous considérons des exemples vérifiant cette hypothèse et nous appliquons les différents tests décrits dans le paragraphe 5. Nous étudions l'influence de la fonction d'observation sur l'efficacité des tests, dans l'exemple 6.2 et celle de la dérive dans l'exemple 6.3. Nous nous intéressons ensuite à un exemple vérifiant l'hypothèse $(HD1)$, présenté dans [4]. Ceci nous permettra d'illustrer que dans ce cas, les décisions sont prises beaucoup plus rapidement que sous $(HD2)$. Nous justifierons ce fait par le comportement de la solution de l'équation nde Zakai.

Dans le paragraphe 5, sous $(HD2)$, nous avons proposé une méthode de détermination des bornes pour les tests du signe et d'estimation des temps d'attente sous l'hypothèse $(H4)$: $B_- < 0$ et $B_+ < 0$. Il est intéressant d'étudier les résultats obtenus dans le

sous l'hypothèse $(H4)$: $B_- < 0$ et $B_+ < 0$. Il est intéressant d'étudier les résultats obtenus dans le cas où l'hypothèse $(H1)$ n'est pas vérifiée. C'est ce que nous nous proposons de faire dans l'exemple 6.6.

Pour chaque exemple, nous commentons les résultats statistiques sur le comportement des tests. On peut consulter ces résultats dans l'annexe A.

Remarque 6.1 On rappelle que la détermination des bornes utilisées dans les tests de décision sur le signe repose sur la possibilité d'atteindre une situation stationnaire. Après la détection d'un passage à zéro, nous devons attendre un "certain" temps, n_0, avant de cumuler les satistiques de décision des tests. Nous avons choisi pour ces exemples $n_0 = 6$.

Exemples sous l'hypothèse (HD2)

Exemple 6.1 On considère le système

$$\begin{cases} x_{k+1} &= x_k + 0.01\, b(x_k) + \sqrt{0.01}\, u_k\,, \quad x_0 \sim \mathcal{N}(-5, 0.1) \\ y_k &= |x_k| + \sqrt{0.01}\, v_k\,, \end{cases}$$

avec

$$b(x) = \begin{cases} -x\,, & x \leq 0 \\ -0.25\, x\,, & x \geq 0\,. \end{cases}$$

On s'intéresse tout d'abord à une trajectoire simulée sur l'intervalle de temps $[0, 10]$.

Dans ce cas, les résultats des différentes procédures étant similaires, nous illustrons par la figure A.1, uniquement ceux des tests de détection et du rapport de vraisemblance sur les observations.

Dans cette figure, la trajectoire du signal à estimer est tracée en trait plein et le filtre proposé en pointillé. Les intervalles de monotonie détectés sont signalés en noir dans le cas où une décision est prise sur le signe et en gris s'il n'y a pas de décision. Nous rappelons que sur un intervalle de monotonie, si le test appliqué décide "signe positif", (respectivement "signe négatif") on trace la sortie du filtre de Kalman-Bucy (2.3) (respectivement la sortie du filtre (2.4)). Si le test ne prend pas de décision, les sorties des deux filtres sont représentées.

Il est intéressant de regarder le comportement de la solution optimale pour ce problème. Dans la figure A.2 on représente, en plus de la trajectoire (en trait plein), l'estimation optimale obtenue par résolution numérique de l'équation de Zakai (en trait pointillé), entourée par sa région de confiance correspondante à une probabilité de 95%. Un générateur automatique de programmes FORTRAN a été utilisé pour obtenir cette solution.

Dans les figures A.3 à A.6, on trace la densité conditionnelle à divers instants de l'intervalle $[0, 10]$. Le trait vertical correspond à la valeur du signal simulé à l'instant

t. On note qu'au moment où la décision sur le signe est prise, la loi conditionnelle approche une loi gaussienne et que près d'un passage à zéro, la loi conditionnelle a deux "bosses".

On peut constater, par exemple, qu'à l'instant $t = 6.5$ les tests ayant déjà détecté un intervalle de monotonie, ne sont pas encore capables de décider sur le signe. Cette décision sera prise aux environs de $t = 8.48$. En regardant le comportement de la solution de l'équation de Zakai, on remarque l'existence de deux "bosses" à partir de l'instant $t = 3$, approximativement. A l'instant $t = 8.48$, instant où les tests de décision sur le signe sont capables de prendre une décision, une des "bosses" de la densité conditionnelle est déjà négligeable (voir la figure ??), ce qui explique le fait qu'une décision puisse être prise. A cet instant, le filtre proposé et l'estimation optimale ont un comportement similaire.

On s'intéresse maintenant aux résultats numériques obtenus pour une trajectoire simulée sur l'intervalle de temps $[0, 100]$ (voir les tableaux A.1 et A.2).

On constate que

- Le test de détection sur le sorties du filtre de Kalman–Bucy permet de détecter des intervalles de monotonie plus longs que le test sur les observations; cependant l'erreur commise est aussi plus grande.

- Le fait d'augmenter la borne de détection (par exemple en diminuant la probabilité d'erreur α) entraine évidemment la détection d'intervalles de monotonie plus petits. Il semble que nous contrôlons moins bien la probabilité d'erreur du test sur la sortie des filtres de Kalman–Bucy que celle du test sur les observations. Voyons par exemple ce qui se passe quand on prend $\alpha_d = 5\%$ (voir le tableau A.2). Le pourcentage d'erreur de détection p_2 est supérieur à celui qui avait été imposé.

- Les résultats des tests de décision sur le signe basés sur le rapport de vraisemblance sur les observations ou sur les sorties du filtre de Kalman sont comparables. Toutes fois, le temps moyen pour prendre une décision semble être légèrement plus petit dans le premier cas que dans le deuxième.

- Le fait de diminuer la probabilité d'erreur pour la décision sur le signe entraine naturellement une réduction du pourcentage d'intervalles avec décision et des temps d'attente plus grands pour prendre ces décisions.

Remarque 6.1 Les temps moyens pour prendre une décision calculés théoriquement sont beaucoup plus élevés que ceux vérifiés en pratique.

Exemple 6.2 On modifie la fonction d'observation dans l'exemple 6.1. On considère:

$$h(x) = 0.4\,|x|\,.$$

On obtient les résultats présentés dans le tableau A.3.

Le fait de diminuer le coefficient $|H_+| = |H_-|$ semble avoir pour effet l'augmentation des pourcentages de détection d'intervalles corrects. Autrement dit, lorsqu'on diminue le coefficient $|H_+| = |H_-|$, les constantes des tests de détection augmentent. On détecte plus de passages à zéro et les longueurs des intervalles de momotonie diminuent. Ce comportement des tests est satisfaisant vu qu'une diminution du coefficient $|H_+| = |H_-|$ est équivalente à une augmentation du bruit d'observation.

On vérifie également une "légère" diminution des temps d'attente pour prendre une décision. Ceci s'explique du fait que les statistiques des tests du signe sont calculés loin des passages à zéro.

Dans ce cas, le test de détection des passages à zéro basé sur les sorties d'un des filtres de Kalman–Bucy est meilleur que celui basé sur les observations.

Exemple 6.3 On considère les mêmes paramètres que dans l'exemple précédent sauf pour la derive $b(x)$:

$$b(x) = \begin{cases} -5\,x\,, & x \leq 0 \\ -0.25\,x\,, & x \geq 0\,. \end{cases}$$

Les résultats sont à consulter dans le tableau A.4.

On remarque que le pourcentage de décisions est maintenant plus élevé et que le temps moyen pour prendre une décision diminue considérablement. Ceci illustre parfaitement les calculs théoriques présentés dans le paragraphe 5. Le fait de diminuer le rapport $|B_+|/(B_+ - B_-)^2$ (respectivement $|B_-|/(B_+ - B_-)^2$) a pour effet d'augmenter le terme de précision dans les statistiques L et $\hat{L}$. On obtient donc de meilleurs résultats, notamment pour le temps d'attente "côté positif" (respectivement "côté négatif").

Exemple sous l'hypothèse (HD1)

Exemple 6.4 Nous changeons la fonction d'observation de l'exemple 6.1:

$$h(x) = \begin{cases} -x\,, & x \leq 0\,, \\ 2\,x\,, & x \geq 0\,. \end{cases}$$

Le système considéré vérifie l'hypothèse $(HD1)$.

Par les figures A.7 à A.10, nous illustrons les résultats des différentes procédures appliquées à une trajectoire simulée sur l'intervalle de temps $[0, 10]$. Nous remarquons que le test de la variation quadratique décide plus souvent que le test sur les sorties des filtres de Kalman-Bucy.

Le comportement de l'estimation optimale et sa région de confiance à 95% représentés dans la figure A.11, montrent clairement que dans ce cas, la loi conditionnelle approche beaucoup plus rapidement une loi gaussienne après un passage à zéro que dans l'exemple 6.1 (voir la figure A.2). Ceci justifie le fait que nous décidons du signe de x_k plus souvent sous l'hypothèse $(HD1)$ que sous $(HD2)$.

La comparaison entre les figures A.12 à A.14 et les figures A.3 à A.6 (voir l'exemple 6.1) illustre à nouveau la remarque précédente. Dans ce cas, une des deux "bosses" de la loi conditionnelle s'estompe rapidement aux environs des instants de décision ($t = 6.86$ et $t = 6.88$), tandis que dans l'exemple 6.1 on ne retrouve un tel comportement qu'aux alentours de l'instant $t = 10$.

Les résultats numériques des tests appliqués à une trajectoire simulée sur l'intervalle de temps $[0, 100]$ (voir le tableau A.5) mettent également en évidence que les temps d'attente pour prendre une décision sous l'hypothèse $(HD1)$ sont nettement plus courts que sous l'hypothèse $(HD2)$.

Remarque 6.5 Quelques remarques sur les résultats des procédures appliquées à plusieurs trajectoires simulées sur un intervalles de temps $[0, 100]$.

- Les tests détectent moins de passages à zéro lorsque l'écart $|H_+ - H_-|$ augmente et le taux d'erreur est supérieur au taux souhaité. De façon plus nette que sous $(HD2)$, nous contrôlons "moins bien" la probabilité d'erreur du test de détection sur la sortie d'un des filtres de Kalman-Bucy que celle du test sur les observations.
- Pour une différence $|\sigma_+ H_+ - \sigma_- H_-| = 0.5$, les tests de détermination du signe ont des taux d'erreur supérieurs à ceux demandés. Le test de la variation quadratique a tendance à décider plus souvent que le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy, mais semble généralement moins fiable. (cf. [4]).
- En augmentant l'écart $|\sigma_+ H_+ - \sigma_- H_-|$, les tests de détermination du signe décident plus souvent et prennent moins de fausses décisions.

Un exemple sous l'hypothèse $(HD2)$ ne vérifiant pas $(H4)$

Exemple 6.6 On considère le système suivant:

$$\begin{cases} x_{k+1} &= x_k + 0.01\, b(x_k) + \sqrt{0.01}\, u_k\,, \quad x_0 \sim \mathcal{N}(-5.0.1) \\ y_k &= |x_k| + \sqrt{0.01}\, v_k\,, \end{cases}$$

avec

$$b(x) = \begin{cases} -x & x \leq 0 \\ 0.2\, x & x \geq 0 \end{cases}$$

Remarque 6.7 Le coefficient de dérive B_+ étant positif et B_- étant négatif, il y a "peu" de passages à zéro et le signal a tendance à "fuir" vers l'infini dès qu'il est du côté positif.

Nous étudions un cas où l' hypothèse $(HD2)$ est vérifiée mais pas $(H4)$. Cependant on applique les tests présentés dans la section 5. On rappelle que les bornes utilisées dans ces tests ont été déterminées sous l'hypothèse $(H4)$ et ne sont plus justifiées dans cet exemple. Il est donc intéressant de regarder les résultats qu'elles engendrent dans une telle situation.

Dans la figure A.15, nous traçons le filtre approché construit à partir du test de détection et du test de rapport de vraisemblance basés sur les observations. Les résultats obtenus par les autres procédures sont similaires.

En comparant le filtre proposé et l'estimation optimale (voir la figure A.16), on remarque que leurs comportements sont semblables sur l'intervalle de monotonie détecté $[0, 3]$ et également à partir de l'instant $t = 7$, environ. Ceci est satisfaisant vu que sur l'intervalle de monotonie $[5.80, 10]$, les instants de décision sur le signe sont $t = 7.15$, pour le test sur les observations, et $t = 7.20$ pour le test sur les sorties des filtres de Kalman-Bucy.

Ainsi, pour certains cas ne vérifiant pas l'hypothèse $(H4)$, on peut encore utiliser les bornes calculées sous cette hypothèse et obtenir des résultats "corrects".

Remarque 6.8 Dans cette situation où l'hypothèse $(H4)$ n'est pas vérifiée, nous ne savons pas estimer théoriquement le temps d'attente d'une décision.

7 Conclusion

On a décrit des tests permettant de déterminer les intervalles de linéarité de la fonction d'observation, pour un système linéaire par morceaux en temps discret avec petit bruit d'observation. Deux types de tests sont nécessaires: un premier test permet de détecter les intervalles où il n'y a pas de passages à zéro; un deuxième test permet de décider du signe pris par la trajectoire sur cet intervalle. Sous l'hypothèse $(HD2)$, différents tests ont été proposés. Pour la détection des intervalles de monotonie, on a le choix entre un test basé sur les observations et un test basé sur la sortie d'un des filtres de Kalman-Bucy. Pour l'étape de décision du signe, on décrit un test de rapport de vraisemblance sur les observations et un test sur les sorties des filtres de Kalman-Bucy.

L'efficacité des tests de détection dépend de la méthode de détermination de leur borne. Pour le test basé sur les observations, une formule permettant d'obtenir cette valeur a été justifiée théoriquement. Ceci n'a pas été de même pour le test basé sur la sortie d'un des filtres de Kalman-Bucy. Il nous semble donc naturel que ce deuxième test se comporte moins bien que le premier. Quant aux tests de détermination du signe, on a vérifié que leurs comportements sont semblables. Des formules permettant de déterminer les bornes et les temps moyens d'attente pour prendre une décision ont été proposées sous l'hypothèse $(H4)$. On remarque de manière générale, que les temps moyens d'attente calculés théoriquement sont plus longs que ceux obtenus en pratique. On note également que l'ordre de grandeur des temps moyens d'attente théoriques et empiriques est plus grand sous l'hypothèse $(HD2)$ que sous $(HD1)$.

On peut envisager d'étendre cette étude à des systèmes où la fonction d'observation h possède plusieurs points critiques dans le cas où l'allure de cette fonction permet d'appliquer les procédures décrites localement. On peut aussi considérer des systèmes de dimension $n \geq 1$, avec des bruits correlés.

Une extension au cas de fonctions monotones par morceaux existe dans [5], sous l'hypothèse $(HD1)$. Des essais ont été faits par [4] dans le cas d'une dynamique non linéaire. Il reste à justifier son analogue sous l'hypothèse $(HD2)$.

Bibliographie

[1] A. Bensoussan : On some approximation techniques in nonlinear filtering, *in Stoch. Diff. Syst., Stoch. Control and Appl.*, W.H. Flemming & P.L. Lions eds., IMA **Vol 10**, Sringer, 1987.

[2] P. Billingsley : Convergence of Probability Measures, John Wiley & Sons, New York, 1968.

[3] W.H. Fleming, D. Ji and E. Pardoux : Piecewise linear filtering with small observation noise, *Proc. 8th Int. Conference on Analysis and Optimization of Systems*, Antibes 1988, Lect. Notes in Control and Info. Sci., n° **111**, Springer, 725-739, 1988.

[4] W.H. Fleming, D. Ji, P. Salame and Q. Zhang : Discrete Time Piecewise Linear Filtering with Small Observation Noise, *Brown University Providence RI 02912*, Division of Applied Mathematics, September 1988 LCDS/CCS 88-27.

[5] W.H. Fleming, E. Pardoux : Piecewise monotone filtering with small observation noise, *Siam J. Control*, **Vol. 27**, No. 5, 1156-1181, Sep. 1989.

[6] W. H.Fleming and R.W.Rishel : Deterministic and Stochastic Optimal Control, Tome 1, Applications of Mathematics, Springer-Verlag, 1975.

[7] D. Ji : Nonlinear filtering with small observation noise, *Ph. D. Thesis*, Brown University, 1987.

[8] R. Katzur, B.Z. Bobrovsky and Z. Schuss : Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel,*SIAM J. Appl. Math.*, **44** Part I: 591-604, Part II: 1176-1191, 1984.

[9] F. Le Gland : Estimation de paramètres dans les processus stochastiques, en observation incomplète - Application à un problème de radio-astronomie, *Thèse de Docteur-Ingenieur*, Université Paris IX-Dauphine, Nov. 1981.

[10] P. Milheiro de Oliveira : Filtres approchés pour un problème de filtrage non linéaire discret avec petit bruit d'observation, *rapport de Recherche INRIA*, **1142**, 1989.

[11] J. Picard : Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio, *SIAM J.Appl. Math.*,**46**, 1098-1125, 1986.

[12] J. Picard : Nonlinear filtering and smoothing with high signal-to-noise ratio, *in Stochastic Processes in Physics and Engeneering*, (Bielefeld 1986), Reidel 1988.

[13] J. Picard : Efficiency of the extended Kalman filter for nonlinear systems with small observation noise, *Rapport de Recherche INRIA*, **1068**, 1989.

Annexe A

Résultats numériques et représentations graphiques des exemples présentés dans la section 6

1 Résultats numériques

Nous introduisons les notations suivantes:

TDO: test de détection sur les observations.
TDK: test de détection sur la sortie d'un des filtres de Kalman Bucy.
TLO: test du rapport de vraisemblance sur les observations.
TLK: test du rapport de vraisemblance sur les sorties des filtres de Kalman -Bucy.
α_d: niveau d'erreur pour le test de détection.
α_s: niveau d'erreur pour le test de détermination du signe.
C: borne du test de détection.
$-l_1$. l_2: intervalle de confiance du test de détermination du signe.
$\mathbf{ET}_-^*$, $\mathbf{ET}_+^*$: temps moyens théoriques d'attente d'une décision.
p_1: pourcentage de pas de temps détectés.
p_2: pourcentage d'intervalles de monotonie corrects.
p_3: pourcentage de pas de temps avec décision.
p_4: pourcentage de décisions correctes.
T_-^*, T_+^*: moyennes empiriques des temps d'attente d'une décision.

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.143	0.143	0.143	0.143
$-l_1\ l_2$	-2.94 2.94	-7.70 7.67	-2.94 2.94	-7.70 7.67
$\mathbf{ET}_-^*\ \mathbf{ET}_+^*$	18.8 4.71	18.9 4.69	18.8 4.71	18.9 4.69
p_1	84.5%	84.5%	86.3%	86.3%
p_2	96.3%	96.3%	93.2%	93.2%
p_3	40.6%	40.6%	41.7%	41.7%
p_4	100%	100%	100%	100%
$\bar{T}_-\ \bar{T}_+$	1.3 1.5	1.3 1.5	2.1 1.6	1.7 1.6

Tableau A.1: résultats numériques pour l'exemple 6.1 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.17	0.17	0.17	0.17
$-l_1\ l_2$	-2.94 2.94	-7.70 7.67	-2.94 2.94	-7.70 7.67
$\mathbf{ET}_-^*\ \mathbf{ET}_+^*$	18.8 4.71	18.9 4.69	18.8 4.71	18.9 4.69
p_1	82.4%	82.4%	84.2%	84.2%
p_2	97.9%	97.9%	97.5%	97.5%
p_3	39.0%	39.0%	41.4%	41.4%
p_4	100%	100%	100%	100%
$\bar{T}_-\ T_+$	1.3 1.5	1.3 1.5	2.1 1.5	1.7 1.5

Tableau A.2: résultats numériques pour l'exemple 6.1 ($\alpha_d = 2.5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.121	0.121	0.302	0.302
$-l_1\ l_2$	-2.94 2.94	-4.38 4.36	-2.94 2.94	-4.38 4.36
$\mathbf{ET}_-^*\ \mathbf{ET}_+^*$	18.8 4.71	18.9 4.70	18.8 4.71	18.9 4.70
p_1	70.8%	70.8%	75.4%	75.4%
p_2	98.4%	98.4%	98.8%	98.8%
p_3	35.0%	34.9%	40.1%	38.8%
p_4	100%	100%	100%	100%
$\bar{T}_-\ \bar{T}_+$	0.0 1.1	0.05 1.2	0.0 1.4	0.05 1.5

Tableau A.3: résultats numériques pour l'exemple 6.2 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.144	0.144	0.144	0.144
$-l_1\ l_2$	-2.94 2.94	-7.69 7.54	-2.94 2.94	-7.69 7.54
$\mathbf{ET}_-^*\ \mathbf{ET}_+^*$	2.350.117	2.400.115	2.350.117	2.400.115
p_1	85.5%	85.5%	87.0%	87.0%
p_2	95.3%	95.3%	92.5%	92.5%
p_3	66.8%	66.8%	69.3%	68.4%
p_4	95.8%	100.0%	90.0%	94.7%
$\bar{T}_-\ \bar{T}_+$	0.0 0.28	0.12 0.29	0.0 0.30	0.12 0.33

Tableau A.4: résultats numériques pour l'exemple 6.3 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection Test du signe	TDO TLO	TDO TLK	TDK TLO	TDK TLK
C	0.165	0.165	0.165	0.165
$-l_1\ l_2$	-4.70 2.37	-15.80 7.80	-4.70 2.37	-15.80 7.80
$\mathbf{E}T_-^*\ \mathbf{E}T_+^*$	0.44 0.13	2.46 0.44	0.44 0.13	2.46 0.44
p_1	86.6%	86.6%	88.2%	88.2%
p_2	93.5%	93.5%	87.5%	87.5%
p_3	71.2%	63.4%	76.1%	65.4%
p_4	82.4%	87.0%	71.1%	66.7%
$\bar{T}_-\ \bar{T}_+$	0.38 0.29	1.60 0.28	0.40 0.37	1.60 0.30

Tableau A.5: résultats numériques pour l'exemple 6.4 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

2 Représentations graphiques

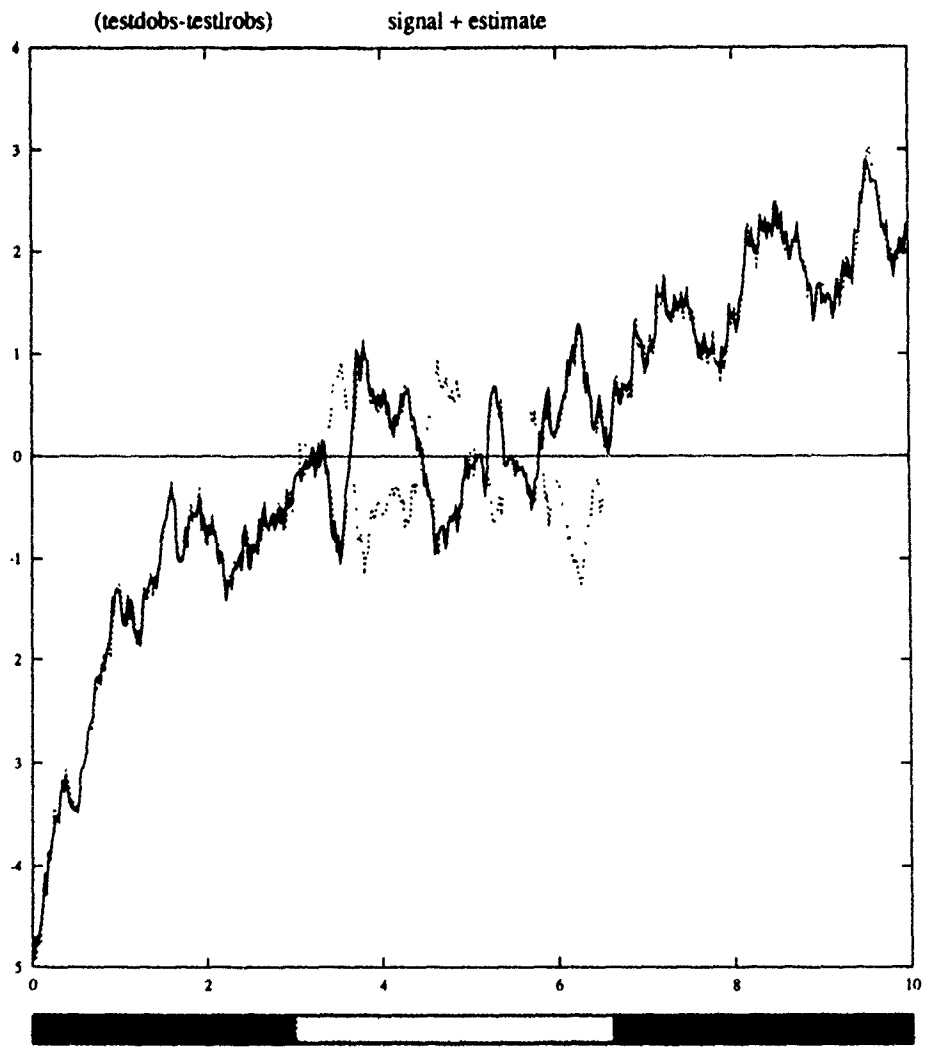

Figure A.1: Sous ($HD2$), exemple 6.1: tests de détection des passages à zero et de décision du signe basés sur les observations.

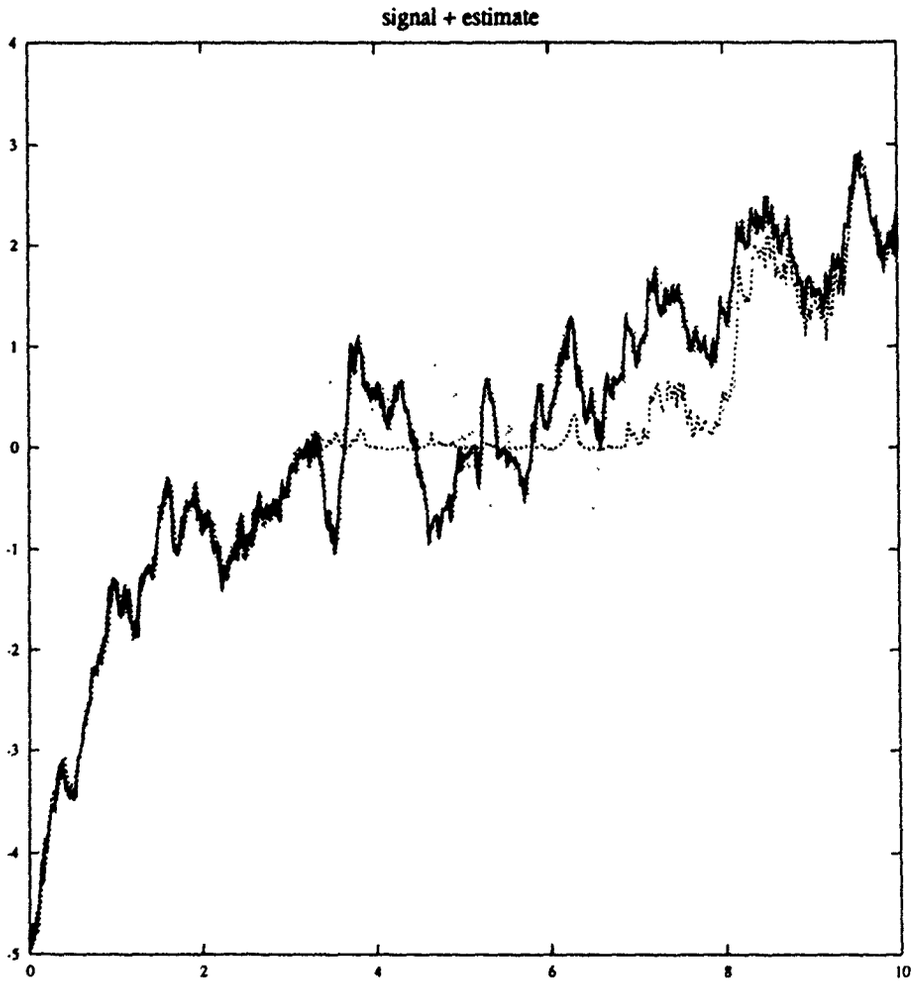

Figure A.2: Sous ($HD2$), exemple 6.1: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

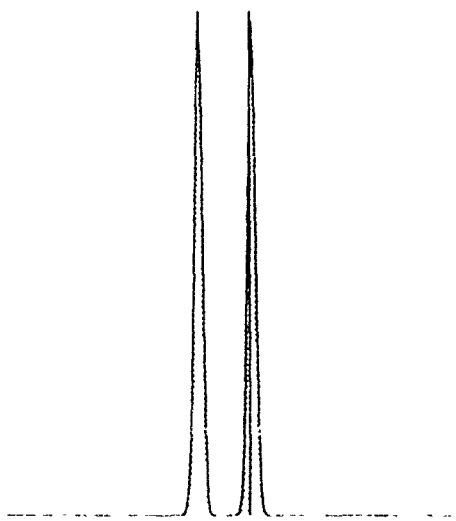

Figure A.3: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 6.5$

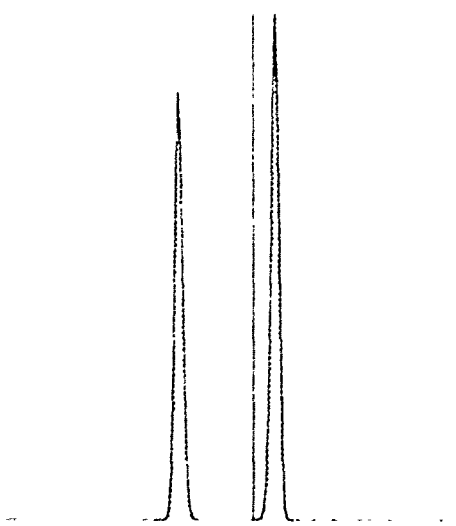

Figure A.4: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 7$

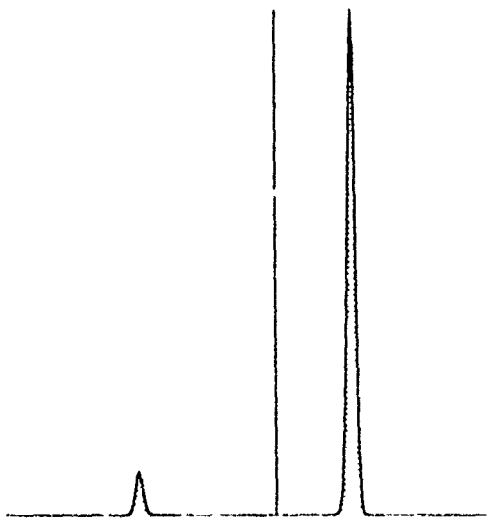

Figure A.5: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 8.48$

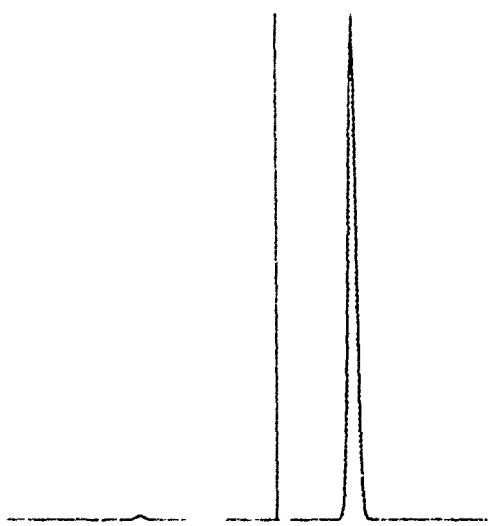

Figure A.6: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 10$

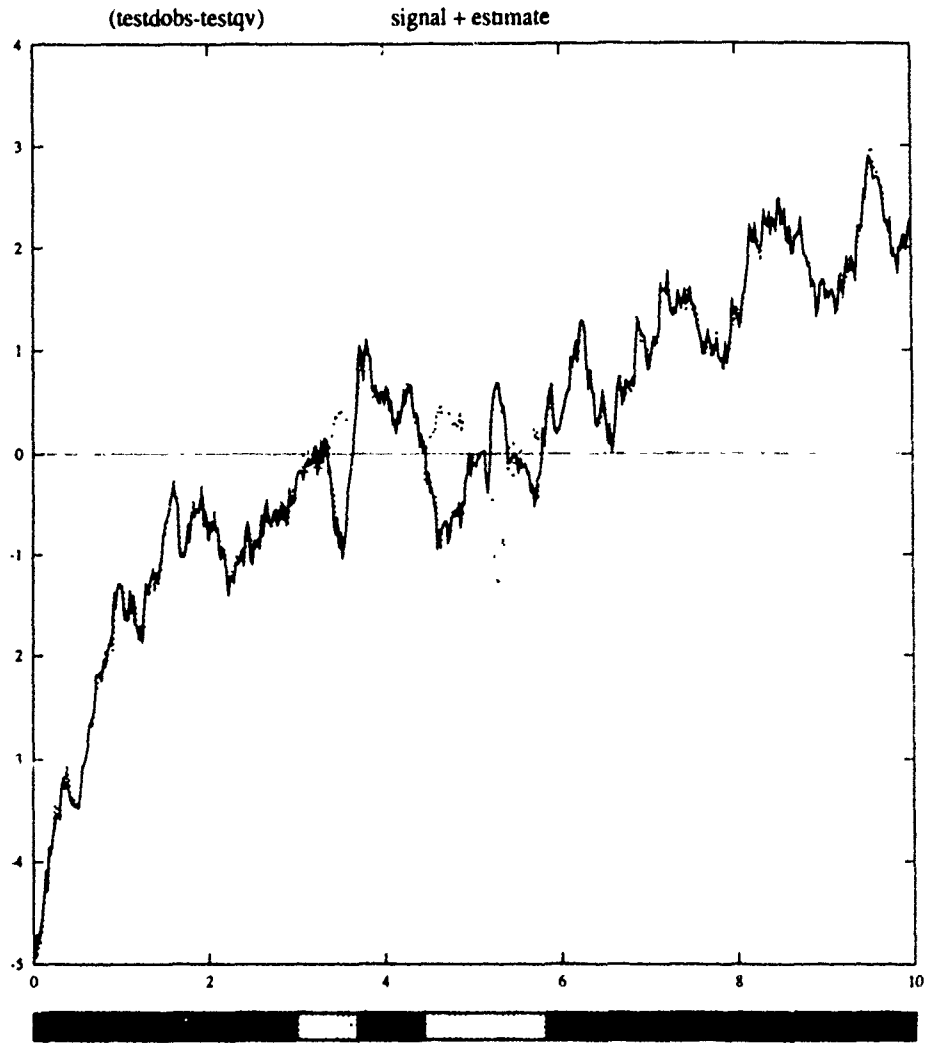

Figure A.7: Sous ($HD1$), exemple 6.1: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la variation quadratique.

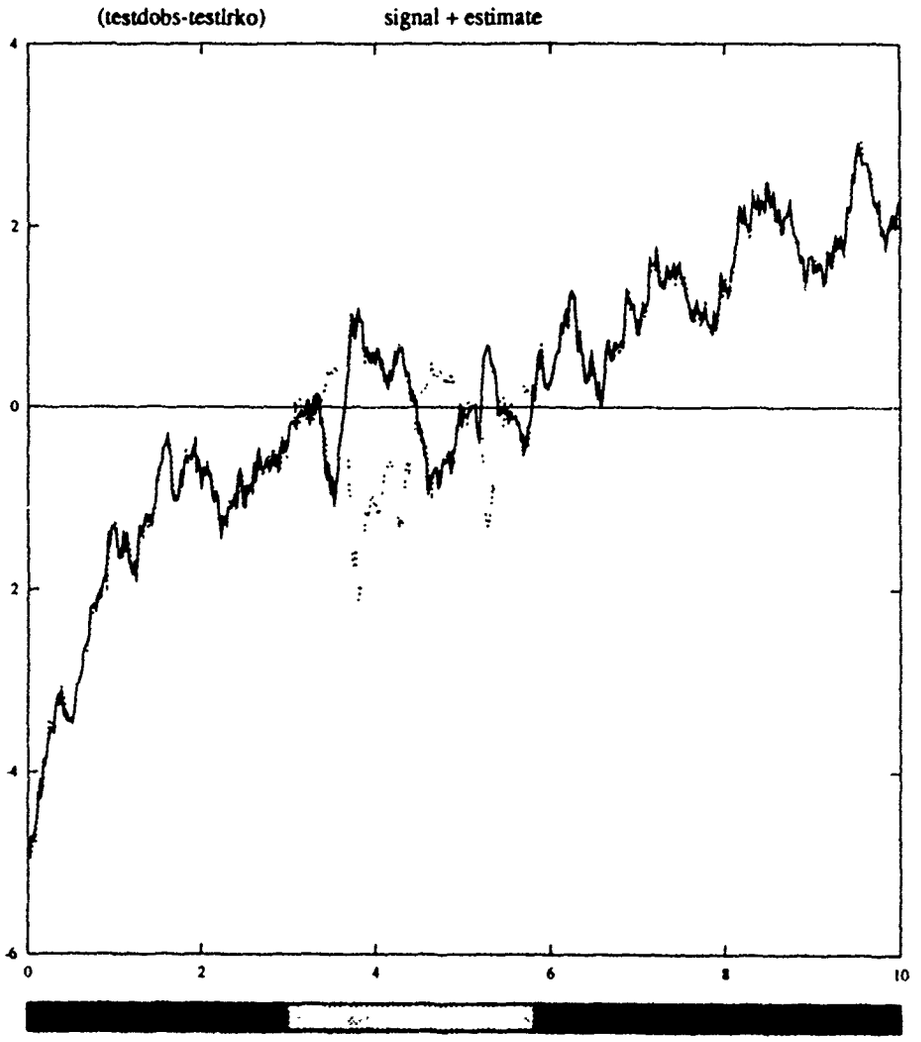

Figure A.8: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la sortie des filtres de Kalman-Bucy.

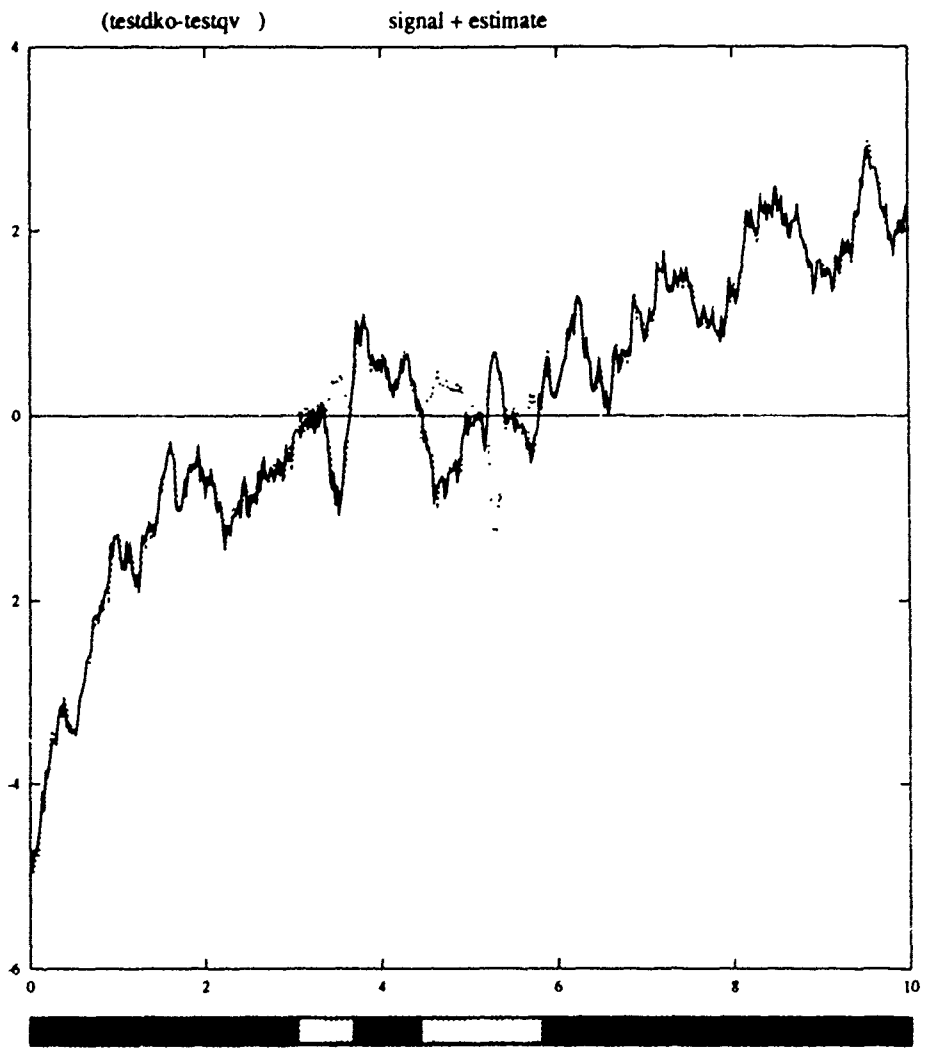

Figure A.9: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur la sortie des filtres de Kalman-Bucy et test de décision du signe basé sur la variation quadratique.

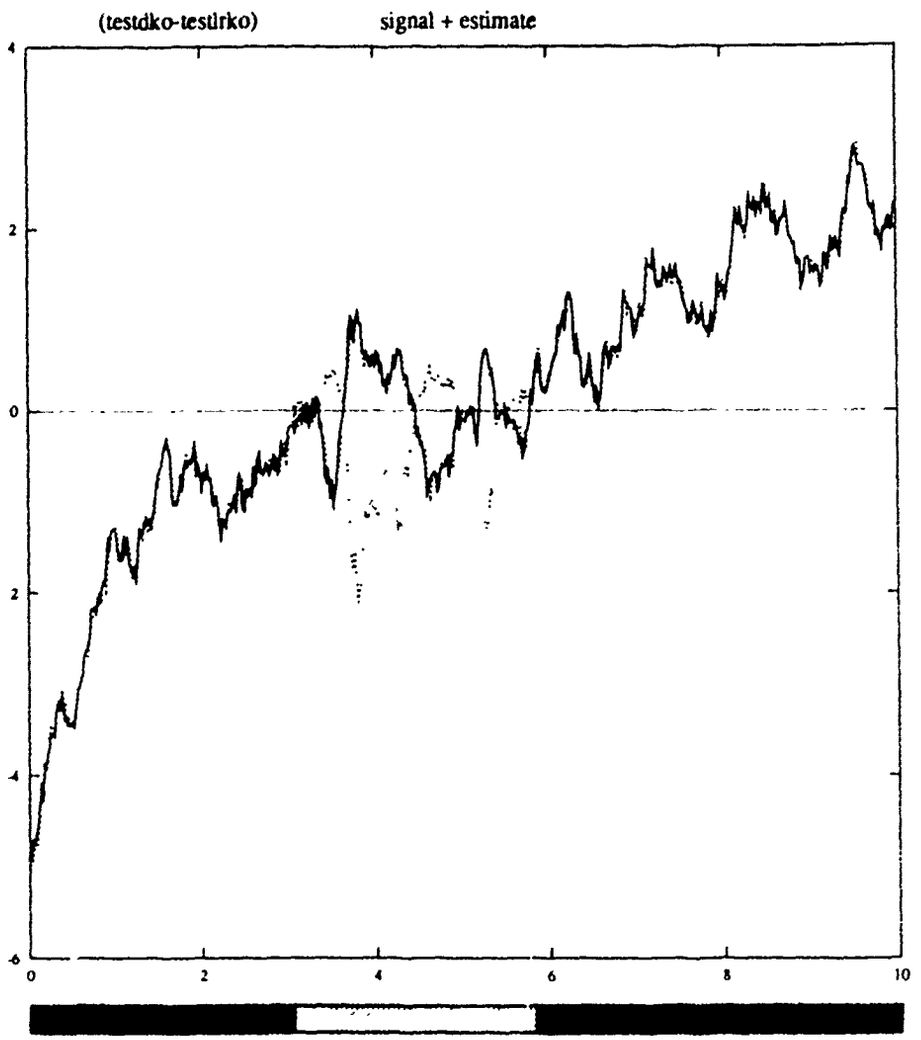

Figure A.10: Sous ($HD1$), exemple 6.4: tests de détection des passages à zero et de décision du signe basés sur la sortie des filtres de Kalman-Bucy.

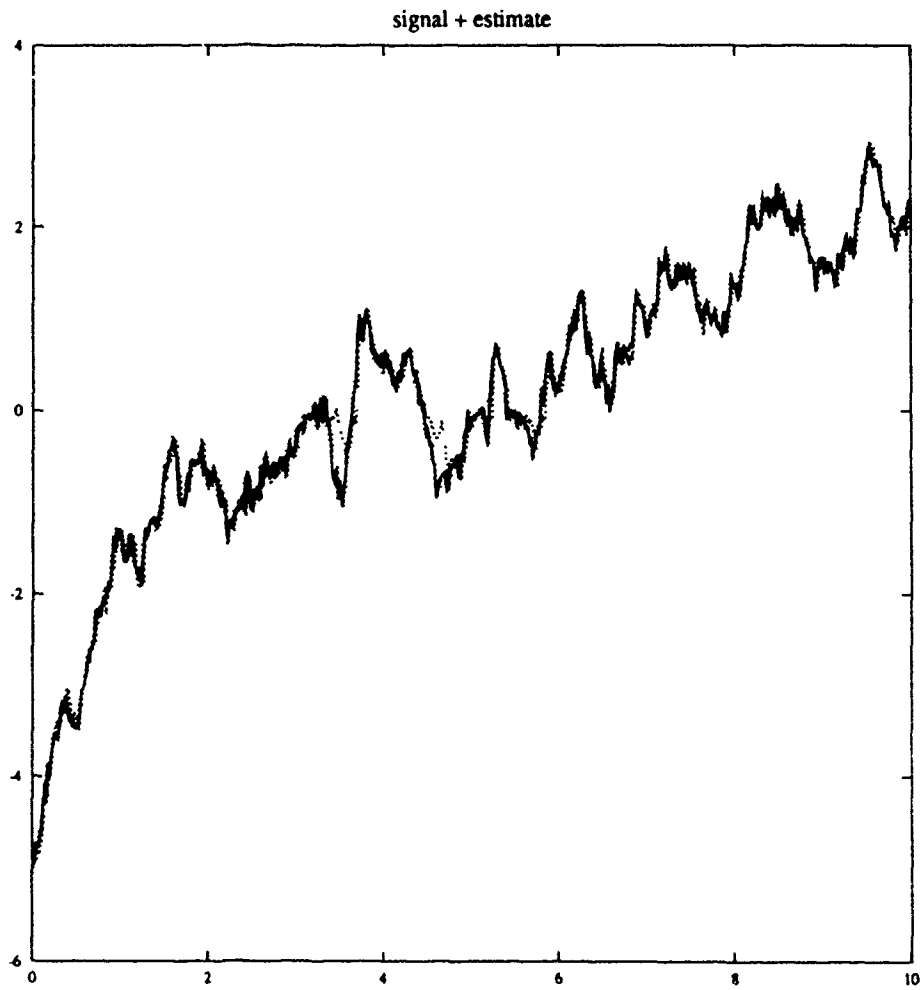

Figure A.11: Sous ($HD1$), exemple 6.4: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

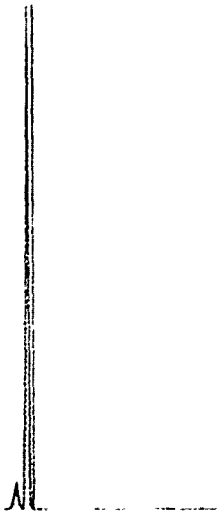

Figure A.12: Sous ($HD1$), exemple 6.1: densité conditionnelle à l'instant $t = 6.60$

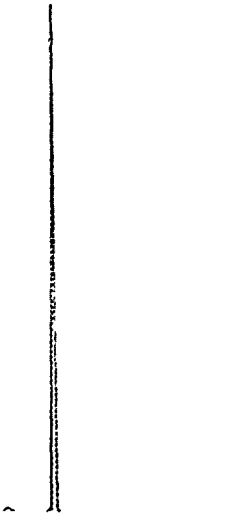

Figure A.13: Sous ($HD1$), exemple 6.4: densité conditionnelle à l'instant $t = 6.86$

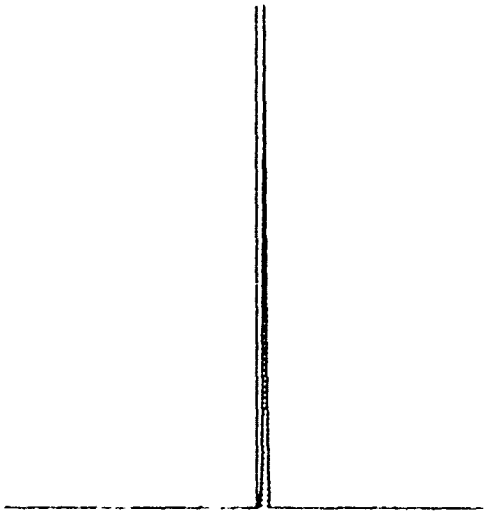

Figure A.14: Sous $(HD1)$, exemple 6.4: densité conditionnelle à l'instant $t = 6.88$

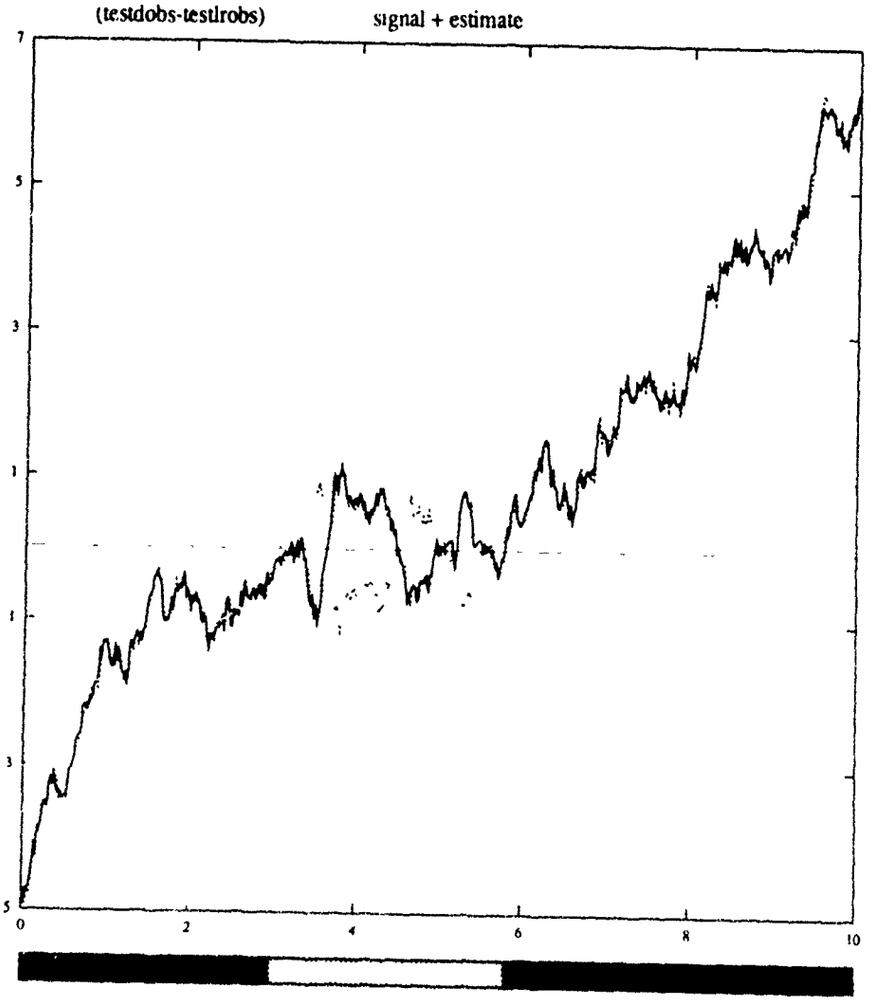

Figure A.15: Exemple 6.6: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la variation quadratique.

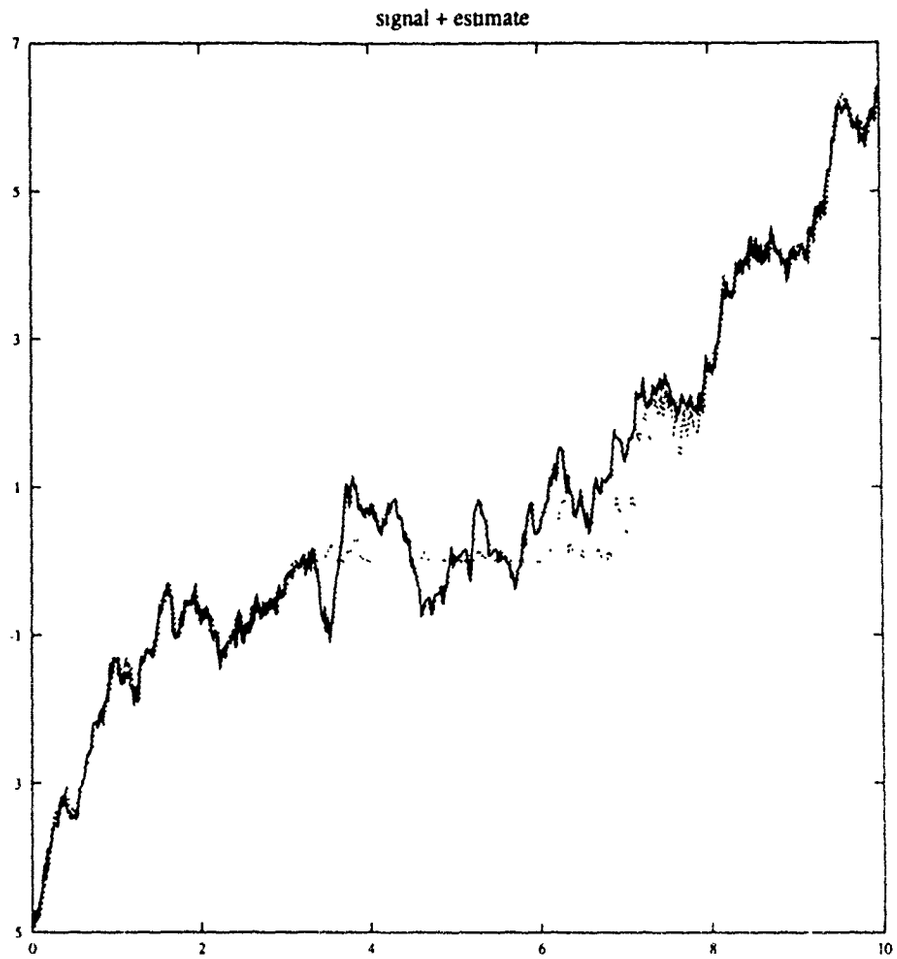

Figure A.16: Exemple 6.6: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%)

Chapitre III

Filtrage linéaire par morceaux partiellement observé avec petit bruit d'observation

Table des matières

III Filtrage linéaire par morceaux partiellement observé avec petit bruit d'observation **III-1**

1 Introduction III-3
2 Le cas $\varepsilon = 0$ III-7
3 Les deux filtres de Kalman-Bucy utilisés pour estimer $X_1(t)$. III-12
4 Détection des passages à zéro de $X_1(t)$ III-18
5 Le test de la variation quadratique III-21
6 Construction du filtre approché III-24
7 Conclusion III-28

Résumé

On considère un problème de filtrage linéaire par morceaux avec petit bruit d'observation dans le cas où le processus d'état de dimension deux est observé par un processus unidimensionel. Sous une certaine "hypothèse de détectabilité", on propose un filtre sous optimal de dimension finie en traitant séparément la composante d'état "directement" mesurée Le filtre approché correspondant à cette composante est construit à l'aide de filtres de Kalman-Bucy et de tests statistiques. A partir de ce dernier, on obtient une estimation pour la deuxième composante. On montre que l'erreur d'approximation est asymptotiquement optimale.

1 Introduction

Nous nous intéressons à une classe de problèmes de filtrage "linéaire par morceaux" partiellement observés avec petit bruit d'observation qui après un changement de variables, peuvent s'écrire sous la forme suivante:

$$\begin{cases} dX_1(t) &= b_{11}(X_1(t))dt + B_{12}X_2(t)dt + \sigma_{11}dV_1(t) + \sigma_{12}dV_2(t) \\ dX_2(t) &= b_{21}(X_1(t))dt + B_{22}X_2(t)dt + \sigma_{21}dV_1(t) + \sigma_{22}dV_2(t) \\ dY_t &= h_1(X_1(t))dt + \varepsilon\, dW(t), \qquad Y_0 = 0\,. \end{cases} \tag{1.1}$$

où $\{X(t) = (X_1(t), X_2(t))^*\,,\ t \geq 0\}$ est le processus d'état à valeurs dans $\mathbb{R}^2$, à estimer à l'instant t, au vu des réalisations jusqu'à l'instant t du processus observé $\{Y(t)\,,\ t \geq 0\}$ à valeurs dans $\mathbb{R}$. Les processus $\{(V_1(t), V_2(t))^*\,,\ t \geq 0\}$ et $\{W(t)\,,\ t \geq 0\}$ sont deux processus de Wiener indépendants à valeurs respectives dans $\mathbb{R}^2$ et $\mathbb{R}$ définis sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ tels que le système (1.1) soit vérifié. De plus, le paramètre ε est petit.

Le problème est linéaire "par morceaux" au sens où les fonctions mesurables b_{11}, b_{21} et h_1 sont définies comme il suit:

$$\begin{aligned} b_{11} : \mathbb{R} &\to \mathbb{R} \\ x &\to b_{11}(x) = \mathbf{1}_{\mathbb{R}_-}(x)B_{11}^- x + \mathbf{1}_{\mathbb{R}_+}(x)B_{11}^+ x, \\ b_{21} : \mathbb{R} &\to \mathbb{R} \\ x &\to b_{21}(x) = \mathbf{1}_{\mathbb{R}_-}(x)B_{21}^- x + \mathbf{1}_{\mathbb{R}_+}(x)B_{21}^+ x, \\ h_1 : \mathbb{R} &\to \mathbb{R} \\ x &\to h_1(x) = \mathbf{1}_{\mathbb{R}_-}(x)H_1^- x + \mathbf{1}_{\mathbb{R}_+}(x)H_1^+ x. \end{aligned}$$

On note h la fonction d'observation, $h : \mathbb{R}^2 \to \mathbb{R}$ est donnée par $h(x_1, x_2) = h_1(x_1)$.

On suppose que:

$(H1)$: $X_1(0) = x_{01} \in \mathbb{R}$, $X_2(0) \sim \mathcal{N}(x_{02}, Q_{02})$ et $X_2(0)$ est indépedant des processus de Wiener $\{V(t) : t \geq 0\}$ et $\{W(t) : t \geq 0\}$.

$(H2)$: h_1 est non injective: $H_1^+ H_1^- < 0$.

$(H3)$: Le système est observable: $B_{12} \neq 0$.

$(H4)$: $\sigma_{11}^2 + \sigma_{12}^2 \neq 0$.

$(H5)$: $\sigma_{21}^2 + \sigma_{22}^2 \neq 0$.

Sur $\mathbb{R}^+ \times \mathbb{R}$ (resp. $\mathbb{R}^- \times \mathbb{R}$), le problème (1.1) s'écrit comme un problème linéaire $(\mathcal{P}_+)$ (resp. $(\mathcal{P}_-)$).

On émet de plus l'hypothèse suivante:

$(H6)$: Les systèmes $(\mathcal{P}_-)$ et $(\mathcal{P}_+)$ sont stabilisables, i.e. il existe une matrice $(2,2)$, L_-, (resp. L_+), telle que la matrice $B_- + \sigma L_-$, (resp. $B_+ + \sigma L_+$), est une matrice stable, avec

$$B_\pm = \begin{pmatrix} B_{11}^\pm & B_{21} \\ B_{12}^\pm & B_{22} \end{pmatrix} \text{ et } \sigma = \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}.$$

Dans le problème linéaire $(\mathcal{P}_+)$, (resp. $(\mathcal{P}_-)$), la composante $X_1(t)$ du processus d'état est directement mesurée. Dans la suite de notre propos, par abus de language, nous la qualifierons également de composante "directement" mesurée pour le problème (1.1).

Notre but est de construire un filtre sous-optimal en dimension finie facilement calculable et d'obtenir des résultats asymptotiques lorsque ε tend vers zéro.

Récemment, plusieurs problèmes de filtrage non linéaire partiellement observés avec petit bruit ou avec petit bruit d'observation ont été étudiés (cf.[11], [6],[14], [9]). Dans [11], J. Picard énonce des conditions sous lesquelles l'erreur de filtrage (i.e. l'erreur quadratique minimale) tend vers zéro lorsque le bruit tend vers zéro et montre que dans ce contexte, le filtre de Kalman étendu est efficace. Cependant, dans le cas où uniquement le bruit d'observation est faible, une des conditions se traduit par l'injectivité forte de la fonction d'observation. Dans [14] et [9], les auteurs s'intéressent à des problèmes particuliers où la dynamique du système est modélisée par un processus de diffusion de dimension 2, le processus d'observation unidimensionnel et la fonction d'observation est injective par rapport à la composante "directement" mesurée. A.J. Krener considère dans [6], un modèle avec petits bruits de dynamique et d'observation pour certaines composantes. La non linéarité du système dépend uniquement de ces composantes qui peuvent être rapidement estimées de façon précise et la fonction d'observation

est supposée linéaire. Par un changement de probabilité, il montre que dans ce type de situation, le filtre linéaire est asymptotiquement optimal.

Dans notre problème, la fonction d'observation est non injective par rapport à $X_1(t)$, mais notre étude s'apparente à celle de A.J. Krener au sens où la non linéarité du système dépend uniquement de la composante $X_1(t)$ que nous pouvons estimer rapidement et de façon précise sous une certaine "hypothèse de détectabilité". En effet, sous l'hypothèse $(HD1)$: $|H_1^-| \neq |H_1^+|$, une procédure de tests semblable à celle décrite dans les paragraphes 3 et 4 du chapitre I, (voir également [1]), permet de déterminer les intervalles de temps durant lesquels $X_1(t)$ ne passe pas par zéro et sur ces intervalles, de décider de son signe avec une probabilité proche de un quand ε est petit.

Sur les intervalles de monotonie de la fonction h_1, le problème (1.1) s'écrit comme un problème linéaire avec condition initiale non gaussienne. Les diverses études sur le comportement asymptotique du filtre de Kalman-Bucy associé à un système linéaire partiellement observé quand le bruit d'observation tend vers zéro (cf. [12], [10], [4], [7]) mettent notamment en évidence que sous les hypothèses émises, la composante du filtre correspondante à $X_1(t)$ est à "mémoire courte", contrairement à celle de $X_2(t)$ non "directement" mesurée.

Ainsi, uniquement la condition initiale de $X_1(t)$ est sans importance et donc sur les intervalles de monotonie de la fonction h_1, uniquement $X_1(t)$ pourra être estimer précisemment par un des deux filtres de Kalman-Bucy (FK_-) ou (FK_+) associés resp. aux systèmes $(\mathcal{P}_-)$ et $(\mathcal{P}_+)$. Dans le cas $\varepsilon = 0$, la différence de comportement des deux composantes du processus d'état est parfaitement illustrée. Nous nous inspirons de cette étude pour proposer sur un intervalle de temps $[0,T]$, sous l'hypothèse $(HD1)$, un filtre approché pour $X_1(t)$ à partir d'une procédure de tests et des deux filtres de Kalman-Bucy (FK_-) et (FK_+). A l'aide de ce dernier, nous construisons un filtre pour $X_2(t)$. On montre que l'erreur d'approximation est asymptotiquement optimale.

Ce chapitre est organisé de la façon suivante:

Dans le paragraphe 2, on étudie le cas $\varepsilon = 0$. Dans le paragraphe 3, on donne quelques résultats préliminaires sur les filtres de Kalman-Bucy (FK_-) et (FK_+). Dans les paragraphes 4 et 5, on présente de façon succincte la procédure de tests permettant de décider du signe de $X_1(t)$ sur les intervalles de temps durant lesquels $X_1(t)$ ne passe pas par zéro et de déterminer ces intervalles avec une probabilité proche de un quand ε est petit. Dans le

paragraphe 6, on propose sur un intervalle de temps $[0,T]$, un filtre sous optimal pour estimer $X(t)$ et on montre que lorsque ε tend vers zéro, l'erreur d'approximation commise tend vers l'erreur quadratique minimale dans le cas $\varepsilon = 0$. On obtient une majoration de la vitesse de convergence. On conclut dans le paragraphe 7.

Notation 1.1 *On notera c ou C, les constantes indépendantes de ε dont les valeurs sont sans importance.*

2 Le cas $\varepsilon = 0$

On considère le modèle suivant:

$$\begin{cases} dX_1(t) &= b_{11}(X_1(t))dt + B_{12}X_2(t)dt + \sigma_{11}dV_1(t) + \sigma_{12}dV_2(t), \\ dX_2(t) &= b_{21}(X_1(t))dt + B_{22}X_2(t)dt + \sigma_{21}dV_1(t) + \sigma_{22}dV_2(t), \\ y_t &= h_1(X_1(t)), \end{cases} \quad (2.1)$$

avec les hypothèses et les définitions établies pour le système (1.1).

Soit $\delta > 0$, on introduit une suite de temps d'arrêt $\{\tau_k\,,\, k \in \mathbb{N}\}$:

$$\begin{aligned} \tau_0 &= 0 \\ \tau_k &= \inf\{t > \tau_{k-1} + \delta \; t.q. \; y(t) = 0\}, \quad k \geq 1, \end{aligned}$$

et on définit les événements:

$$\begin{aligned} A^0_-(\tau_k + \delta, \tau_{k+1}) &= \{X_1(t) \leq 0\,;\, \tau_k + \delta \leq t \leq \tau_{k+1}\} \\ A^0_+(\tau_k + \delta, \tau_{k+1}) &= \{X_1(t) \geq 0\,;\, \tau_k + \delta \leq t \leq \tau_{k+1}\}. \end{aligned}$$

Sur $A^0_-(\tau_k + \delta, \tau_{k+1})$:

$$\langle y \rangle_t - \langle y \rangle_{\tau_k + \delta} = (H_1^-)^2(\sigma_{11}^2 + \sigma_{12}^2)\,(t - \tau_k - \delta).$$

Sur $A^0_+(\tau_k + \delta, \tau_{k+1})$:

$$\langle y \rangle_t - \langle y \rangle_{\tau_k + \delta} = (H_1^+)^2(\sigma_{11}^2 + \sigma_{12}^2)\,(t - \tau_k - \delta).$$

On pose

$$\begin{aligned} \Gamma_-^2 &= (H_1^-)^2(\sigma_{11}^2 + \sigma_{12}^2), \\ \Gamma_+^2 &= (H_1^+)^2(\sigma_{11}^2 + \sigma_{12}^2). \end{aligned}$$

Ainsi, sous l'hypothèse $(HD1)$: $\left|H_1^+\right| \neq \left|H_1^-\right|$, pour $t \in [\tau_k + \delta, \tau_{k+1}]$, la variation quadratique de $y(t)$ permet de décider du signe de $X_1(t)$.

Soit $m_0 = \left[\dfrac{T}{\delta}\right] + 1$, sur l'intervalle $[0, T]$, $X_1(t)$ peut être estimé par le processus $\{\bar{X}_1^0(t)\,;\, 0 \leq t \leq T\}$ défini de la manière suivante:

$$\begin{aligned} \bar{X}_1^0(t) &= \sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k + \delta, \tau_{k+1}]}(t) \Big[\mathbf{1}_{\{\langle y \rangle_t - \langle y \rangle_{\tau_k + \delta} = \Gamma_-^2 (t - \tau_k - \delta)\}}\, y_t / H_1^- \\ &\qquad + \mathbf{1}_{\{\langle y \rangle_t - \langle y \rangle_{\tau_k + \delta} = \Gamma_+^2 (t - \tau_k - \delta)\}}\, y_t / H_1^+ \Big], \\ \bar{X}_1^0(0) &= x_{01}. \end{aligned} \quad (2.2)$$

On remarque que pour $t \in [\tau_k + \delta, \tau_{k+1}]$, $X_1(t)$ est calculé de façon exacte. Il n'en est pas de même pour la composante $X_2(t)$.

Dans un premier temps, introduisons le processus $\{M_2^0(t)\,;\, t \geq 0\}$ avec $M_2^0(t) = \mathbf{E}(X_2(t)|\mathcal{X}_t^1)$ et $\mathcal{X}_t^1 = \sigma(X_1(s)\,;\, 0 \leq s \leq t)$. Si nous considérons le processus $\{X_1(t)\,;\, t \geq 0\}$ comme l'observation, on obtient un nouveau problème de filtrage:

$$\left\{\begin{array}{rcl} dX_2(t) &=& B_{22}X_2(t)\,dt + b_{21}(X_1(t))\,dt + \sigma_{21}dV_1(t) + \sigma_{22}dV_2(t), \\ dX_1(t) &=& B_{12}X_2(t)\,dt + b_{11}(X_1(t))dt + \sigma_{11}dV_1(t) + \sigma_{12}dV_2(t), \end{array}\right. \tag{2.3}$$

avec comme condition initiale $X_1(0) = x_{01} \in \mathbb{R}$ et $X_2(0) \sim \mathcal{N}(x_{02}, Q_{02})$. On pose

$$\begin{array}{rcl} f_1^2 &=& \sigma_{11}^2 + \sigma_{12}^2, \\ f_2^2 &=& \sigma_{21}^2 + \sigma_{22}^2, \\ f_{12} &=& \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}. \end{array}$$

D'après U. Haussmann et E. Pardoux [3], (voir aussi R.CH. Liptzer, A.N. Shyriaev [8]), les hypothèses émises et le lemme 2.7 chapitre I, la loi conditionnelle est gaussienne. Elle est notée $\mathcal{N}(M_2^0(t), R_{22}^0(t))$ où $\{(M_2^0(t), R_{22}^0(t))\,;\, t \geq 0\}$ est l'unique solution forte du système différentiel

$$\left\{\begin{array}{rcl} dM_2^0(t) &=& [B_{22}M_2^0(t) + b_{21}(X_1(t))]\,dt \\ && +K^0(t)\,[dX_1(t) - (b_{11}(X_1(t)) + B_{12}M_2^0(t))\,dt], \\ M_2^0(0) &=& x_{02}, \\ && \text{avec} \quad K^0(t) = \frac{1}{f_1^2}(R_{22}^0(t)B_{12} + f_{12}), \\ \dot{R}_{22}^0(t) &=& 2\left(B_{22} - B_{12}\frac{f_{12}}{f_1^2}\right)R_{22}^0(t) - \frac{(B_{12}R_{22}^0)^2}{f_1^2} + f_2^2 - \frac{f_{12}^2}{f_1^2}. \\ R_{22}^0(0) &=& Q_{02}. \end{array}\right. \tag{2.4}$$

Remarque 2.1 Ce filtre a la même "allure" qu'un filtre de Kalman-Bucy.

Remarque 2.2 Si nous considérons un nouveau processus d'état du filtre défini par

$$\tilde{M}_2^0(t) \triangleq M_2^0(t) - K^0(t)X_1(t),$$

nous évitons la différentiation du processus $\{X_1(t)\,;\, t \geq 0\}$. (Cf. [10]).

Le système (2.3) étant supposé observable ($(H3)$: $B_{12} \neq 0$) et commandable ($(H5)$: $f_2^2 > 0$), on a:

$$\lim_{t \to +\infty} R_{22}^0(t) = R_{22}^0 > 0,$$

où R_{22}^0 est solution de l'équation algébrique de Riccati associé à (2.4).

$$R_{22}^0 = \frac{f_1^2}{B_{12}^2}\left(B_{22} - \frac{B_{12}f_{12}}{f_1^2} + \sqrt{\Delta}\right), \tag{2.5}$$

$$\Delta = \left(B_{22} - \frac{B_{12}f_{12}}{f_1^2}\right)^2 + \frac{B_{12}^2}{f_1^2}\left(f_2^2 - \frac{f_{12}^2}{f_1^2}\right) > 0. \tag{2.6}$$

On obtient une expression explicite pour la matrice de covariance $R_{22}^0(t)$:

$$R_{22}^0(t) = R_{22}^0 + \frac{2e^{-2t\sqrt{\Delta}} f_{11}^2 \sqrt{\Delta}(Q_{02} - R_{22}^0)}{2f_{11}^2\sqrt{\Delta} + B_{12}^2(Q_{02} - R_{22}^0)(1 - e^{-2t\sqrt{\Delta}})}. \tag{2.7}$$

On s'inspire de ce qui précède pour proposer un filtre approché, $\bar{X}_2^0(t)$ au filtre optimal de $X_2(t)$ que nous noterons $\hat{X}_2^0(t)$. D'après la remarque 2.2, on a

$$\bar{X}_2^0(t) = \tilde{X}_2^0(t) + K^0(t)\bar{X}_1^0(t), \tag{2.8}$$

où $K^0(t)$ est donné en (2.4) et le processus $\{\tilde{X}_2^0(t)\,;\, t \geq 0\}$ est défini par

$$\begin{cases} \frac{d}{dt}\tilde{X}_2^0(t) &= [B_{22} - K^0(t)B_{12}]\,\tilde{X}_2^0(t) + b_{21}(\bar{X}_1^0(t)) \\ & \quad -K^0(t)b_{11}(\bar{X}_1^0(t)) - \dot{K}^0(t)\bar{X}_1^0(t), \\ \tilde{X}_2^0(0) &= x_{02}. \end{cases} \tag{2.9}$$

Remarque 2.3 Etant donné que $\sigma(y_t\,;\, t \geq 0) \subset \mathcal{X}_t^1$, on a

$$\mathbf{E}|M_2^0(t) - X_2(t)|^2 \leq \mathbf{E}|\hat{X}_2^0(t) - X_2(t)|^2.$$

De l'égalité $\mathbf{E}|M_2^0(t) - X_2(t)|^2 = R_{22}^0(t)$, on déduit l'ordre de grandeur suivant

$$\mathbf{E}|\hat{X}_2^0(t) - X_2(t)|^2 = \mathcal{O}(1).$$

Estimons les erreurs d'approximation.

Théorème 2.4 *$\forall T > 0$, il existe $\delta_0 > 0$, $c > 0$, tels que $\forall \delta \in (0, \delta_0]$,*

$$\mathbf{E}|\bar{X}_1^0(t) - X_1(t)|^2 \leq c\,\delta, \tag{2.10}$$

$$\mathbf{E}|\bar{X}_2^0(t) - M_2^0(t)|^2 \leq c\,\delta, \tag{2.11}$$

et

$$\mathbf{E}|\bar{X}_2^0(t) - X_2(t)|^2 \leq \mathbf{E}|\hat{X}_2^0(t) - X_2(t)|^2 + c\,\delta^{1/2}. \tag{2.12}$$

Preuve Considérons la différence $X_1(t) - \bar{X}_1^0(t)$.

$$X_1(t) - \bar{X}_1^0(t) = \sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t) X_1(t).$$

Etant donné que $X_1(\tau_k) = 0$, pour $t \geq \tau_k$, $X_1(t)$ s'écrit:

$$\begin{aligned} X_1(t) &= \int_{\tau_k}^t b_{11}(X_1(s))ds + \int_{\tau_k}^t B_{12}(X_2(s))ds \\ &\quad + \sigma_{11}(\,V_1(t) - V_1(\tau_k)\,) + \sigma_{12}(\,V_2(t) - V_2(\tau_k)\,). \end{aligned}$$

$$\begin{aligned} &\mathbf{E}\left|\sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t) X_1(t)\right|^2 \leq \mathbf{E}\left[\sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t)|X_1(t)|^2\right] \\ \leq\ & c\,\delta^2 \sup_{[0,T]} |X_1(t)|^2 + c\,\mathbf{E}\left[\sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}|V_1(t) - V_1(\tau_k)|^2\right] \\ &+ c\,\mathbf{E}\left[\sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t)|V_2(t) - V_2(\tau_k)|^2\right]. \end{aligned}$$

Vu que $\mathbf{E}\left(\sup_{[0,T]} |X(t)|^2\right) < C(T)$, il suffit de traiter le second ou le troisième terme. En conditionnant par $\mathcal{F}_{\tau_k}$ et en utilisant l'indépendance de $V_1(t \wedge (\tau_k + \delta)) - V_1(\tau_k)$ avec cette tribu, on obtient:

$$\begin{aligned} &\mathbf{E}\left[\sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t)|V_1(t) - V_1(\tau_k)|^2\right] \\ \leq\ & \mathbf{E}\left[\sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t)|V_1(t \wedge (\tau_k + \delta)) - V_1(\tau_k)|^2\right] \\ \leq\ & \mathbf{E}\left[\sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t)(\,t \wedge (\tau_k + \delta) - \tau_k\,)\right] \ \leq\ \delta. \end{aligned}$$

On en déduit la majoration suivante:

$$\mathbf{E}|\bar{X}_1^0(t) - X_1(t)|^2 \leq c\,\delta + c(T)\,\delta^2 \tag{2.13}$$

Estimons $\mathbf{E}|\bar{X}_2^0(t) - M_2^0(t)|^2$.

On pose $\beta_t = -B_{22} + K^0(t)B_{12}$. D'après les expressions (2.4) et (2.7), il existe $t_0 > 0$ tel que $\forall t > t_0$, $\beta_t > 0$ et $K^0(t)$ est borné uniformement en t.

$$\mathbf{E}|\bar{X}_2^0(t) - M_2^0(t)|^2 \leq 2\,\mathbf{E}|\bar{X}_2^0(t) - \tilde{M}_2^0(t)|^2 + 2\,c\mathbf{E}|X_1(t) - \bar{X}_1^0(t)|^2.$$

D'après (2.13), il suffit de considérer le premier terme.

$$\begin{aligned}\bar{X}_2^0(t) - \tilde{M}_2^0(t) = & \int_0^t e^{-\int_s^t \beta_r dr}[b_{21}(\bar{X}_1^0(s)) - b_{21}(X_1(s))]\,ds \\ & - \int_0^t e^{-\int_s^t \beta_r dr}[b_{11}(\bar{X}_1^0(s)) - b_{11}(X_1(s))]\,ds \\ & - \int_0^t e^{-\int_s^t \beta_r dr}\dot{K}(s)[\bar{X}_1^0(s)) - b_{11}(X_1(s)]\,ds.\end{aligned}$$

Des inégalités suivantes

$$|b_{21}(\bar{X}_1^0(s)) - b_{21}(X_1(s))|^2 \leq \sup(|B_{21}^-|, |B_{21}^+|)^2 \sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t)|X_1(s)|^2,$$

$$|b_{11}(\bar{X}_1^0(s)) - b_{11}(X_1(s))|^2 \leq \sup(|B_{11}^-|, |B_{11}^+|)^2 \sum_{k=0}^{m_0} \mathbf{1}_{[\tau_k, \tau_k+\delta[}(t)|X_1(s)|^2$$

et de la majoration (2.13), on obtient

$$\mathbf{E}|\bar{X}_2^0(t) - \tilde{M}_2^0(t)|^2 \leq c\,\delta + c(T)\,\delta^2. \tag{2.14}$$

On en déduit (2.11).

Comparons maintenant $\mathbf{E}|\bar{X}_2^0(t) - X_2(t)|^2$ à l'erreur de filtrage i.e. à $\mathbf{E}|\hat{X}_2^0(t) - X_2(t)|^2$. Si on note $\|.\|_2$ la norme dans L^2, on obtient (2.12) en utilisant l'inégalité triangulaire

$$\|\bar{X}_2^0(t) - X_2(t)\|_2 \leq \|\bar{X}_2^0(t) - M_2^0(t)\|_2 + \|M_2^0(t) - X_2(t)\|_2\,,$$

la majoration (2.14) et la remarque 2.3.

□

3 Les deux filtres de Kalman–Bucy utilisés pour estimer $X_1(t)$

Sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, P^+)$, on considère le problème linéaire $(\mathcal{P}_+)$ suivant:

$$\begin{cases} dX(t) &= B_+X(t)dt + \sigma dV^+(t), \\ dY(t) &= H_+X(t)dt + \varepsilon\, dW^+(t), \end{cases} \tag{3.1}$$

où $\{V^+(t);\, t \geq 0\}$ et $\{W^+(t);\, t \geq 0\}$ sont deux processus de Wiener standards à valeurs respectives dans $\mathbb{R}^2$ et $\mathbb{R}$, sous la loi P^+. Les matrices $(2,2)$, B_+ et σ et la matrice $(1,2)$, H_+ sont définies par

$$B_+ = \begin{pmatrix} B_{11}^+ & B_{12} \\ B_{21}^+ & B_{22} \end{pmatrix} \qquad \sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

et

$$H_+ = (H_1^+, 0).$$

On suppose que les hypothèses $(H1)$ à $(H6)$ sont vérifiées.

Construisons le filtre de Kalman-Bucy $(\hat{X}^+(t), R_+^\varepsilon(t))$, correspondant au système (3.1). D'après les hypothèses $(H3)$ et $(H6)$, le système est observable et stabilisable donc

$$\lim_{t \to +\infty} R_+^\varepsilon(t) = R_+^\varepsilon$$

et la matrice $B_+ - \frac{1}{\varepsilon^2} R_+^\varepsilon (H_1^+)^2$ est stable. (Cf. [13]).

Prennons pour loi initiale du filtre $\mathcal{N}(x_0, R_+^\varepsilon)$ avec $x_0 = (x_{01}, x_{02})^*$ et R_+^ε, la matrice de covariance conditionnelle stationnaire

$$R_+^\varepsilon = \begin{pmatrix} R_{11}^{\varepsilon,+} & R_{12}^{\varepsilon,+} \\ R_{12}^{\varepsilon,+} & R_{22}^{\varepsilon,+} \end{pmatrix},$$

où $R_{11}^{\varepsilon,+} \geq 0$, $R_{22}^{\varepsilon,+} \geq 0$ et $R_{12}^{\varepsilon,+}$ sont solutions des équations suivantes:

$$\begin{cases} 0 &= 2B_{11}^+ R_{11}^{\varepsilon,+} + 2B_{12} R_{12}^{\varepsilon,+} + f_1^2 - \frac{1}{\varepsilon^2}(H_1^+ R_{11}^{\varepsilon,+})^2, \\ 0 &= (B_{11}^+ + B_{22}) R_{12}^{\varepsilon,+} + B_{12} R_{22}^{\varepsilon,+} + B_{21}^+ R_{11}^{\varepsilon,+} + f_{12} \\ &\qquad\qquad - \frac{1}{\varepsilon^2}(H_1^+)^2 R_{11}^{\varepsilon,+} R_{12}^{\varepsilon,+}, \\ 0 &= 2B_{21}^+ R_{12}^{\varepsilon,+} + 2B_{22} R_{22}^{\varepsilon,+} + f_2^2 - \frac{1}{\varepsilon^2}(H_1^+ R_{12}^{\varepsilon,+})^2. \end{cases} \tag{3.2}$$

On obtient le lemme suivant:

Lemme 3.1 *Sous les hypothèses (H2) à (H6), pour ε suffisament petit, on a:*

1)
$$R_{11}^{\varepsilon,+} = \varepsilon R_{11}^{+} + \bar{R}_{11}^{\varepsilon,+} \tag{3.3}$$

avec
$$R_{11}^{+} = \frac{\sqrt{f_1^2}}{|H_1^{+}|} > 0 \quad et \quad \bar{R}_{11}^{\varepsilon,+} \leq \mathcal{O}(\varepsilon^2). \tag{3.4}$$

2)
$$R_{12}^{\varepsilon,+} = \varepsilon R_{12}^{+} + \bar{R}_{12}^{\varepsilon,+}, \tag{3.5}$$

avec
$$R_{12}^{+} = \frac{B_{12}R_{22}^{+} + f}{|H_1^{+}|\sqrt{f_1^2}} \quad et \quad \bar{R}_{12}^{\varepsilon,+} \leq \mathcal{O}(\varepsilon^2). \tag{3.6}$$

3)
$$R_{22}^{\varepsilon,+} = R_{22}^{+} + \bar{R}_{22}^{\varepsilon,+} \tag{3.7}$$

avec
$$R_{22}^{+} = \frac{f_1^2}{B_{12}^2}\left[B_{22}^{+} - \frac{B_{12}f_{12}}{f_1^2} + \sqrt{\Delta}\right] > 0, \tag{3.8}$$
$$\Delta = \left[B_{22}^{+} - \frac{B_{12}f_{12}}{f_1^2}\right]^2 + \frac{B_{12}^2}{f_1^2}\left[f_2^2 - \frac{f_{12}^2}{f_1^2}\right] > 0$$
$$\bar{R}_{22}^{\varepsilon,+} \leq \mathcal{O}(\varepsilon).$$

Preuve D'après le système (3.2), sous les hypothèses $(H2)$ et $(H3)$, il existe $\varepsilon_0 > 0$, r_1, r_2, $r_{12} > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$R_{11}^{\varepsilon,+} \leq \varepsilon r_1, \quad R_2^{\varepsilon,+} \leq r_2 \text{ et } |R_{12}^{\varepsilon,+}| \leq \varepsilon r_{12}.$$

Donc, il existe une sous-suite de $\frac{R_{11}^{\varepsilon,+}}{\varepsilon}$, $R_{22}^{\varepsilon,+}$ et $\frac{R_{12}^{\varepsilon,+}}{\varepsilon}$ qui converge quand $\varepsilon \to 0$. Or, toutes sous-suites convergentes de $\frac{R_{11}^{\varepsilon,+}}{\varepsilon}$, $R_{22}^{\varepsilon,+}$ et $\frac{R_{12}^{\varepsilon,+}}{\varepsilon}$ ont respectivement pour limite $R_{11}^{+} \geq 0$, $R_{22}^{+} \geq 0$ et R_{12}^{+} solutions du système suivant:

$$\begin{cases} 0 &= f_1^2 - (H_1^{+}R_{11}^{+})^2, \\ 0 &= B_{12}R_{22}^{+} + f_{12} - (H_1^{+})^2 R_{11}^{+}R_{12}^{+}, \\ 0 &= 2B_{22}R_{22}^{+} + f_2^2 - (H_1^{+}R_{12}^{+})^2. \end{cases}$$

De la résolution de ce système sous les hypothèses (H_2)-$(H5)$, on déduit les résultats énoncés dans le lemme.

□

Remarque 3.2 La matrice R_{22}^+ est égale à la matrice de covariance de $X_2(\cdot)$ sachant $\mathcal{X}_t^1$ dans le cas où $\varepsilon = 0$.

Soit $\hat{X}^+(t)$, l'espérance conditionnelle de $X(t)$ sachant $\mathcal{Y}_t$, le processus $\{\hat{X}^+(t) = (\hat{X}_1^+(t), \hat{X}_2^+(t))^* ; t \geq 0\}$ est solution de l'équation différentielle stochastique suivante:

$$\left\{\begin{array}{rcl} d\hat{X}_1^+(t) & = & \left(B_{11}^+ - \frac{1}{\varepsilon}K_1^+ H_1^+\right)\hat{X}_1^+(t)dt + B_{12}\hat{X}_2^+(t)dt + \frac{1}{\varepsilon}K_1^+ h_1(X_1(t))dt \\ & & \qquad + K_1^+ dW(t), \\ \hat{X}_1^+(0) & = & x_{01}, \\ d\hat{X}_2^+(t) & = & B_{22}\hat{X}_2^+(t)dt + \left(B_{21}^+ - \frac{1}{\varepsilon}K_2^+ H_1^+\right)\hat{X}_1^+(t)dt + \frac{1}{\varepsilon}K_2^+(t)h_1(X_1(t))dt \\ & & \qquad + K_2^+ dW(t), \\ \hat{X}_2(0) & = & x_{02}, \end{array}\right. \tag{3.9}$$

où le processus $\{W(t) ; t \geq 0\}$ est un processus de Wiener standard sous la loi P et la matrice K_+ est égale à ε fois la matrice gain,

$$K_+ = \begin{pmatrix} K_1^+ \\ K_2^+ \end{pmatrix} = \frac{1}{\varepsilon}\begin{pmatrix} R_{11}^{\varepsilon,+} H_1^+ \\ R_{12}^{\varepsilon,+} H_1^+ \end{pmatrix}.$$

Ainsi, d'après le lemme 3.1, on a les ordres de grandeur:

$$K_1^+ = \mathcal{O}(1), \tag{3.10}$$

$$K_2^+ \leq \mathcal{O}(1). \tag{3.11}$$

On montre le résultat préliminaire suivant:

Lemme 3.3 *$\forall T > 0$, il existe $c > 0$, $C > 0$ tels que*

$$\limsup_{\varepsilon \to 0} \mathbf{E}\left(\exp\left\{c \sup_{[0,T]} |\hat{X}^+(t)|^2\right\}\right) < C.$$

Remarque 3.4 On précise que $\mathbf{E}$ désigne l'espérance par rapport à la loi P.

Preuve

$$\hat{X}^+(t) = e^{\Lambda_+^\varepsilon t}x_0 + \frac{1}{\varepsilon}\int_0^t e^{\Lambda_+^\varepsilon(t-s)}K_+h(X_s)ds + \int_0^t e^{\Lambda_+^\varepsilon(t-s)}K_+dW(s),$$

avec

$$\Lambda_+^\varepsilon = \begin{pmatrix} B_{11}^+ - \frac{1}{\varepsilon}K_1^+H_1^+ & B_{12} \\ B_{21}^+ - \frac{1}{\varepsilon}K_2^+H_1^+ & B_{22} \end{pmatrix}$$

Soit ϕ_t la solution de l'équation matricielle :

$$\dot{\phi}_t = \Lambda_+^\varepsilon \phi_t\,, \quad \phi_0 = I_2.$$

$$\phi_t = e^{\Lambda_+^\varepsilon t}.$$

On vérifie que les valeurs propres de la matrice Λ_+^ε sont d'ordre $\mathcal{O}(1)$ et $\mathcal{O}(\frac{1}{\varepsilon})$. Ainsi, la stabilité de la matrice implique: il existe $\varepsilon_0 > 0$, $\lambda > 0$ et $C > 0$ tels que $\forall 0 < \varepsilon \leq \varepsilon_0$,

$$|e^{\Lambda_+^\varepsilon(t-s)}| \leq Ce^{-\lambda(t-s)} \quad 0 \leq s \leq t\,.$$

Sachant par le lemme 2.7 chapitre I que pour tout $T > 0$ il existe $c > 0$, $C > 0$ tels que

$$\mathbf{E}\left(\exp\left\{c\sup_{[0,T]}|X(t)|^2\right\}\right) < C, \tag{3.12}$$

on montre facilement qu'il existe $\varepsilon_0 > 0$, $c > 0$, $C > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$,

$$\mathbf{E}\left(\exp\left\{c\varepsilon^2\sup_{[0,T]}|\hat{X}^+(t)|^2\right\}\right) < C. \tag{3.13}$$

On en déduit le résultat intermédiare suivant:

$$\mathbf{E}\left(\exp\left\{c\varepsilon^2\sup_{[0,T]}|\hat{X}_2^+(t)|^2\right\}\right) < C. \tag{3.14}$$

Etablissons la majoration suivante:

$$\limsup_{\varepsilon \to 0}\mathbf{E}\left(\exp\left\{c\sup_{[0,T]}|\hat{X}_1^+(t)|^2\right\}\right) < C. \tag{3.15}$$

$$\begin{aligned}\hat{X}_1^+(t) &= e^{-\frac{1}{\epsilon}\lambda_1(t-s)}x_{01} + B_{22}\int_e^t e^{-\frac{1}{\epsilon}\lambda_1(t-s)}\hat{X}_2^+(s)ds \\ &\quad +\frac{1}{\varepsilon}\int_e^t e^{-\frac{1}{\epsilon}\lambda_1(t-s)}K_1^+ h_1(X_1(s))ds \\ &\quad +\int_e^t e^{-\frac{1}{\epsilon}\lambda_1(t-s)}K_1^+ dW(s), \end{aligned} \tag{3.16}$$

avec $\lambda_1 = -\varepsilon B_{11}^+ + K_1^+ H_1^+ > 0$ pour ε suffisamant petit, d'après (3.10). On a donc l'inégalité

$$\begin{aligned}\sup_{[0,T]}|\hat{X}_1^+(t)|^2 &\leq 4|x_{01}|^2 + 4\varepsilon^2|B_{22}|^2 \sup_{[0,T]}|\hat{X}_2^+(t)|^2 \\ &\quad +4k\sup_{[0,T]}|X_1(t)|^2 + 4\sup_{[0,T]}\left|\int_0^t e^{-\frac{1}{\epsilon}\lambda_1(t-s)}K_1^+ dW(s)\right|.\end{aligned}$$

Le premier terme est déterministe. On traite le second et le troisième terme respectivement à l'aide de (3.14) et (3.12). Par un raisonnement similaire à celui effectué dans la démonstration du lemme 2.8 chapitre I, on montre qu'il existe $c > 0$, $C > 0$ tels que

$$\limsup_{\epsilon \to 0}\mathbf{E}\left(\exp\left\{c\sup_{[0,T]}\left|\int_0^t e^{-\frac{1}{\epsilon}\lambda_1(t-s)}K_1^+ dW(s)\right|^2\right\}\right) < C.$$

On en déduit (3.15).

Pour démontrer le lemme, il nous reste donc à établir le même type de majoration pour $\hat{X}_2^+(t)$. En posant $Z_t = h_1(X_1(t)) - H_1^+\hat{X}_1^+(t)$, on obtient les équations

$$d\hat{X}_1^+(t) = B_{11}^+\hat{X}_1^+(t)dt + B_{12}\hat{X}_2^+(t)dt + \frac{1}{\varepsilon}K_1^+ Z_t dt + K_1^+ dW(t), \tag{3.17}$$

$$d\hat{X}_2^+(t) = B_{22}\hat{X}_2^+(t)dt + B_{21}^+\hat{X}_1^+(t)dt + \frac{1}{\varepsilon}K_2^+ Z_t dt + K_2^+ dW(t). \tag{3.18}$$

D'après (3.4), $K_1^+ \neq 0$, et on a l'égalité suivante:

$$\frac{1}{\varepsilon}Z_t dt = \frac{1}{K_1^+}\left(d\hat{X}_1^+(t) - B_{11}^+\hat{X}_1^+(t)dt - B_{12}\hat{X}_2^+(t)dt - K_1^+ dW(t)\right). \tag{3.19}$$

En portant (3.19) dans l'équation (3.18), on obtient

$$\begin{aligned} d\hat{X}_2^+(t) &= \left(B_{22} - \frac{K_2^+}{K_1^+}B_{12}\right)\hat{X}_2^+(t)dt \\ &\quad + \left(B_{12} - \frac{K_2^+}{K_1^+}B_{11}^+\right)\hat{X}_1^+(t)dt + \frac{K_2^+}{K_1^+}d\hat{X}_1^+(t). \end{aligned}$$

On note $\beta = -B_{22} + \frac{K_2^+}{K_1^+}B_{12}$. Les résultats du lemme 3.3 impliquent: $\beta > 0$, pour ε suffisamment petit.

$$\begin{aligned} \hat{X}_2^+(t) &= e^{-\beta(t-s)}x_{02} + \left(B_{12} - \frac{K_2^+}{K_1^+}B_{11}^+\right)\int_\epsilon^t e^{-\beta(t-s)}\hat{X}_1^+(s)ds \\ &\quad + \frac{K_2^+}{K_1^+}\int_\epsilon^t e^{-\beta(t-s)}d\hat{X}_1^+(s). \end{aligned} \tag{3.20}$$

D'après (3.15), il suffit de s'intéresser au troisième terme. Par intégration par parties, on obtient:

$$\begin{aligned} \int_\epsilon^t e^{-\beta(t-s)}d\hat{X}_1^+(s) &= \hat{X}_1^+(t) - x_{01}e^{-\beta t} \\ &\quad + \beta\int_\epsilon^t e^{-\beta(t-s)}\hat{X}_1^+(s)ds. \end{aligned}$$

On déduit donc, de la majoration (3.15) l'existence de $c > 0$ et $C > 0$ tels que

$$\limsup_{\epsilon \to 0} \mathbf{E}\left(\exp\left\{c\sup_{[0,T]}|\hat{X}_2^+(t)|^2\right\}\right) < C.$$

□

Les resultats que nous avons démontrés dans ce paragraphe pour le filtre de Kalman-Bucy (FK_+) s'adaptent sans problème au filtre (FK_-).

Remarque 3.5 Les équations (3.16) et (3.20) mettent en évidence que le filtre $\hat{X}_1(t)$ correspondant à la composante $X_1(t)$ a une "mémoire" décroissante quand ε tend vers zéro contrairement à celui de la composante $X_2(t)$ non "directement" observée. Par la suite, les filtres Kalman-Bucy (FK_-) et (FK_+) n'étant utilisés que pour construire un filtre approché pour $X_1(t)$, la condition initiale du filtre est sans importance. Ceci justifie le choix d'une loi initiale avec variance conditionnelle stationnaire.

4 Détection des passages à zéro de $X_1(t)$

On s'intéresse premièrement aux écarts $|h_1(X_1(t)) - H_1^- \hat{X}_1^-(t)|$ et $|h_1(X_1(t)) - H_1^+ \hat{X}_1^+(t)|$.

Proposition 4.1 $\forall \theta > 0$, $\forall \alpha \in [0, \frac{1}{2})$, $\forall \gamma \in (2\alpha, 1)$, $\forall \mu \in (0,1)$, *il existe* $k > 0$, $\varepsilon_0 > 0$ *tel que* $\forall \varepsilon \in (0, \varepsilon_0]$,

$$P\left(\left\{\sup_{[\varepsilon^\mu, T]} |h_1(X_1(t)) - H_1^- \hat{X}_1^-(t)| > \theta\varepsilon^\alpha\right\}\right) \leq e^{-k/\varepsilon^{\gamma-2\alpha}},$$

$$P\left(\left\{\sup_{[\varepsilon^\mu, T]} |h_1(X_1(t)) - H_1^+ \hat{X}_1^+(t)| > \theta\varepsilon^\alpha\right\}\right) \leq e^{-k/\varepsilon^{\gamma-2\alpha}}.$$

Remarque 4.2 la durée ε^μ avec $\mu \in (0,1)$ représente le temps nécessaire pour que la différence initiale n'est plus d'influence.

Preuve Les deux majorations se démontrant de façon similaire, on considère uniquement la seconde. D'aprés la formule d'Itô–Tanaka [5], on a:

$$dh_1(X_1(t)) = h_1'(X_1(t))dX_1(t) + \frac{1}{2}(H_1^+ - H_1^-)dL_t^0(X_1).$$

où

- h_1' est la dérivée à gauche de h_1.

$$h_1'(x) = H_1^- \mathbf{1}_{\{X_1 \leq 0\}} + H_1^+ \mathbf{1}_{\{X_1 > 0\}},$$

- Le processus $\{L_t^0\,;\, t \geq 0\}$ est le temps local de la semi-martingale X_1.

D'autre part,

$$\begin{aligned} dH_1^+ \hat{X}_1^+(t) &= H_1^+ B_{11}^+ \hat{X}_1^+(t)\,dt + H_1^+ B_{12} \hat{X}_2^+(t)\,dt \\ &\quad + \frac{1}{\varepsilon} H_1^+ K_1^+ [dY_t - H_1^+ \hat{X}_1^+(t)dt]. \end{aligned}$$

On pose $Z_t = h_1(X_1(t)) - H_1^+ \hat{X}_1^+(t)$. On obtient

$$\begin{aligned} Z_t &= e^{-\frac{1}{\varepsilon}\lambda t} Z_0 + \int_0^t e^{-\frac{1}{\varepsilon}\lambda(t-s)} \varphi_s ds \\ &\quad + \int_0^t e^{-\frac{1}{\varepsilon}\lambda(t-s)} (\sigma_{11} dV_1(s) + \sigma_{12} dV_2(s)) \\ &\quad + \lambda \int_0^t e^{-\frac{1}{\varepsilon}\lambda(t-s)} dW_s \\ &\quad + \frac{1}{2}(H_1^+ - H_1^-) \int_0^t e^{-\frac{1}{\varepsilon}\lambda(t-s)} dL_s^0, \end{aligned} \tag{4.1}$$

avec

- $\lambda = H_1^+ K_1^+ > 0$,
- $\varphi_t = h_1'(X_1(t))\,[b_{11}(X_1(t)) + B_{12} X_2(t)] - H_1^+ B_{11}^+ \hat{X}_1^+(t) - H_1^+ B_{12} \hat{X}_2^+(t)$.

On considère les termes de droite de l'égalité (4.1) séparément.

$$\sup_{[\varepsilon^\mu, T]} |e^{-\frac{1}{\varepsilon}\lambda t} Z_0| \leq \exp\left\{-\frac{\lambda}{\varepsilon^{1-\mu}}\right\} |Z_0|.$$

Vu que Z_0 est déterministe, pour ε suffisament petit, on a

$$\exp\left\{-\frac{\lambda}{\varepsilon^{1-\mu}}\right\} |Z_0| \leq \frac{\theta \varepsilon^\alpha}{5}.$$

En utilisant le lemme 3.3, on traite les autres termes par un raisonnement similaire à celui de la démonstration de la proposition 3.1 chapitre I, et on obtient le résultat demandé.

□

Soit $c_a > c_b > 0$, $0 \leq \alpha < \frac{1}{2}$ et $0 < \mu < 1$ choisis, on introduit les temps d'arrêts suivants:

$$\begin{aligned} \tau_0^b &= 0, \\ \tau_k^a &= \inf\{t > \tau_{k-1}^b \ t.q. \ |\hat{X}_1^+(t)| \geq c_a \varepsilon^\alpha\} \wedge T, \quad k > 0, \\ \tau_k^b &= \inf\{t > \tau_k^a \ t.q. \ |\hat{X}_1^+(t)| \leq c_b \varepsilon^\alpha\} \wedge T \quad k > 0. \end{aligned}$$

On définit les événements:

$$\begin{aligned} A_+^k &= A_+(\tau_k^a, \tau_k^b) = \{X_1(t) > 0\,;\, \tau_k^a \leq t \leq \tau_k^b\}, \\ A_-^k &= A_-(\tau_k^a, \tau_k^b) = \{X_1(t) < 0\,;\, \tau_k^a \leq t \leq \tau_k^b\}. \end{aligned}$$

Théorème 4.3 $\forall \alpha \in [0, \frac{1}{2})$, $\forall \gamma \in (2\alpha, 1)$, *il existe* $\varepsilon_0 > 0$, $c > 0$ *tels que* $\forall \varepsilon \in (0, \varepsilon_0]$, $\forall k \in \mathbb{N}^*$,

$$P\left((A_+^k \cup A_-^k)^c\right) \leq e^{-c/\varepsilon^{\gamma-2\alpha}}.$$

Preuve Il existe $\theta > 0$ tel que

$$(A_+^k \cup A_-^k)^c \subset \left\{ \sup_{[0,T]} |h_1(X_1(t)) - H_1^+ \hat{X}_1^+(t)| > \theta \varepsilon^\alpha \right\}.$$

Le résultat découle alors de la proposition 4.1.

□

5 Le test de la variation quadratique

Nous supposons que l'hypothèse $(HD1)$: $|H_1^+| \neq |H_1^-|$ est vérifiée. On représente par $\delta\varepsilon^\mu$ avec $\delta > 0$ et $0 < \mu < 1$, la durée minimale d'un intervalle de monotonie $[\tau_k^a, \tau_k^b]$ pour qu'une décision puisse être prise avec une probabilité d'erreur donnée. Les valeurs de δ et μ sont à déterminer en fonction de cette probabilité d'erreur.

Soit $m = [\delta\varepsilon^{\mu-1}] + 1$. Pour $k \in \mathbb{N}^*$, $j = 0, \ldots, m$, on introduit les notations suivantes:

$$\begin{aligned}
\tau_j &= \tau_k^a + j\varepsilon \\
\bar{y}_j^k &= \frac{1}{\varepsilon}(Y(\tau_{j+1}) - Y(\tau_j)) \\
s_j^k &= \frac{1}{\varepsilon}\int_{\tau_j}^{\tau_{j+1}} h_1(X_1(s))\,ds, \\
\bar{w}_j^k &= W(\tau_{j+1}) - W(\tau_j).
\end{aligned}$$

On note que $\bar{y}_j^k = s_j^k + \bar{w}_j^k$.

Remarque 5.1 $\mathbf{E}(\bar{w}_j^k)^2 = \varepsilon$.

On pose

$$Z_k = \frac{1}{\delta\varepsilon^\mu} \sum_{j=0, j\ pair}^{m-2} (\bar{y}_{j+1}^k - \bar{y}_j^k)^2.$$

Sur $A_+^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}$:

$$Z_k = \frac{1}{\delta\varepsilon^\mu} \sum_{j=0, j\ pair}^{m-2} (\alpha_j^k + \beta_j^k)^2$$

avec

$$\begin{aligned}
\alpha_j^k &= \frac{1}{\varepsilon} H_1^+ \int_{\tau_j}^{\tau_{j+1}} [\sigma_{11}(V_1(s+\varepsilon) - V_1(s)) + \sigma_{12}(V_2(s+\varepsilon) - V_2(s))]ds \\
&\qquad + \bar{w}_{j+1}^k - \bar{w}_j^k, \\
\beta_j^k &= \frac{1}{\varepsilon} H_1^+ \int_{\tau_j}^{\tau_{j+1}} \int_s^{s+\varepsilon} (B_{11}^+ X_1(u) + B_{12}^+ X_2(u))\,du\,ds.
\end{aligned}$$

On note que le vecteur aléatoire α_j^k suit une loi normale $N(0, 2\varepsilon(1+\frac{1}{3}\Gamma_+^2))$ avec $\Gamma_+^2 = (H_1^+)^2(\sigma_{11}^2+\sigma_{12}^2)$ et que la sous suite $\{\alpha_j^k, j\ pair\}$ est une sous suite de variables aléatoires à valeurs dans $\mathbb{R}$, indépendantes et identiquement distribuées.

Quand $\epsilon \to 0$, montrons les convergences suivantes en probabilité:

Sur $A_-^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}$:

$$Z_k \to 1 + \frac{1}{3}\Gamma_-^2.$$

Sur $A_+^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}$:

$$Z_k \to 1 + \frac{1}{3}\Gamma_+^2.$$

Ainsi sur $\{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}$, sous l'hypothèse $(HD1)$, il sera possible de choisir le signe de $X_1(t)$ pour $\tau_k^a \le t \le \tau_k^b$ au vu de Z_k.

Lemme 5.2 *Soit $\delta > 0$, $\mu \in (0,1)$, $\forall\theta > 0$, il existe $\varepsilon_0 > 0$ et $c > 0$ tels que $\forall\varepsilon \in (0,\varepsilon_0]$, $\forall k \in \mathbb{N}^*$,*

$$P\left(\left\{|Z_k - (1+\frac{1}{3}\Gamma_+^2)| \ge \theta\right\} \cap A_+^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}\right) \le e^{-c/\varepsilon^\beta},$$

$$P\left(\left\{|Z_k - (1+\frac{1}{3}\Gamma_-^2)| \ge \theta\right\} \cap A_-^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}\right) \le e^{-c/\varepsilon^\beta},$$

avec $\beta = \frac{1}{2} \wedge (1-\mu)$ □

Preuve La démonstration est similaire à celle du lemme 4.1 chapitre I (voir également [1] [lemme 4.1]).

Construisons le test de la variation quadratique:

On suppose que $(H_1^+)^2 > (H_1^-)^2$. On définit les événements tests suivants:

$$Q_-^k = Q_-^\varepsilon(\tau_k^a, \tau_k^b) = \left\{Z_k < \left(1 + \frac{\Gamma_+^2 + \Gamma_-^2}{6}\right)\right\}$$

$$Q_+^k = Q_+^\varepsilon(\tau_k^a, \tau_k^b) = \left\{Z_k \ge \left(1 + \frac{\Gamma_+^2 + \Gamma_-^2}{6}\right)\right\}$$

Théorème 5.3 *Soit $\delta > 0$, $\mu \in (0,1)$, $\alpha \in [0, \frac{1}{2})$ et $\gamma \in (2\alpha, 1)$, il existe $\varepsilon_0 > 0$ et $c > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$, $\forall k \in \mathbb{N}^*$,*

$$\begin{aligned} P(A_-(\tau_k^a, \tau_k^b)^c \cap Q_-^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}) &\leq e^{-c/\varepsilon^\beta}, \\ P(A_+(\tau_k^a, \tau_k^b)^c \cap Q_+^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}) &\leq e^{-c/\varepsilon^\beta}, \end{aligned}$$

avec $\beta = \frac{1}{2} \wedge (1-\mu) \wedge (\gamma - 2\alpha)$.

Preuve

$$\begin{aligned} (A_+^k)^c \cap Q_+^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\} \subset & \quad (A_-^k \cup A_+^k)^c \\ & \cup \; A_-^k \cap Q_+^k \cap \{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\} \end{aligned}$$

Donc, d'après le théorème 4.3 et le lemme 5.2, on déduit le résultat.

□

6 Construction du filtre approché

On s'inspire du cas $\varepsilon = 0$ pour construire un filtre approché du filtre optimal. Notre but est d'obtenir des résultats asymptotiques. On définit le processus $\{\bar{X}_1(t)\,;\, 0 \le t \le T\}$ de la manière suivante:

$$\begin{aligned} \bar{X}_1(t) &= \sum_{k=1}^{\infty} \mathbf{1}_{[\tau_k^a, \tau_k^b]}(t) \mathbf{1}_{\{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}} \left[\mathbf{1}_{Q_-^k} \hat{X}_1^-(t) + \mathbf{1}_{Q_+^k} \hat{X}_1^+(t) \right], \\ \bar{X}_1(0) &= x_{01}. \end{aligned} \tag{6.1}$$

On note que le filtre proposé est non adapté et discontinu.

Remarque 6.1 Si $t \in [\tau_k^a, \tau_k^b]$, $|\hat{X}_1^+(t)| > c_b \varepsilon^\alpha$.

Remarque 6.2 Sur $t \in [\tau_{k-1}^b, \tau_k^a]$, $X_1(t)$ est proche de zéro avec une probabilité proche de 1 pour ε petit. On l'estime donc par zéro.

D'autre part, sur $[\tau_k^a, \tau_k^a + \delta\varepsilon^\mu]$, $X_1(t)$ est proche de $c_a\, \varepsilon^\alpha$ avec un probabilité proche de 1 quand ε est petit. Nous l'approchons également par zéro.

En utilisant la remarque 2.2, on pose

$$\bar{X}_2(t) = \tilde{X}_2(t) + K^0(t)\bar{X}_1(t), \tag{6.2}$$

avec $K^0(t)$ donné en (2.4) et le processus $\{\tilde{X}_2(t)\,;\, t \ge 0\}$ défini par

$$\begin{aligned} \frac{d}{dt}\tilde{X}_2(t) &= \left[B_{22} - K^0(t)B_{12}\right] \tilde{X}_2(t) + b_{21}(\bar{X}_1(t)) \\ &\quad - K^0(t) b_{11}(\bar{X}_1(t)) - \dot{K}^0(t)\bar{X}_1(t), \\ \tilde{X}_2(0) &= x_{02}. \end{aligned} \tag{6.3}$$

Montrons que lorsque ε tend vers zéro, l'erreur d'estimation tend vers l'erreur de filtrage correspondante au cas $\varepsilon = 0$.

Nous rappelons que les valeurs α et μ interviennent respectivement dans le test de détection des passages à zéro de $X_1(t)$ et le test de la variation quadratique et sont à déterminer pour une probabilité d'erreur donnée.

Théorème 6.3 *Soit $\alpha \in [0, \frac{1}{2})$, $\mu \in (0,1)$, $\forall T > \varepsilon^\mu$, il existe $\varepsilon_0 > 0$, $c > 0$ tels que $\forall \varepsilon \in (0, \varepsilon_0]$, $\forall t \in [\varepsilon^\mu, T]$,*

$$\mathbf{E}\left|\bar{X}_1(t) - X_1(t)\right|^2 \leq c\varepsilon^{2(\alpha\wedge\mu)}, \tag{6.4}$$

$$\mathbf{E}\left|\bar{X}_2(t) - M_2^0(t)\right|^2 \leq c\varepsilon^{2(\alpha\wedge\mu)}, \tag{6.5}$$

$$\mathbf{E}\left|\bar{X}_2(t) - X_2(t)\right|^2 \leq \mathbf{E}\left|\hat{X}_2^0(t) - X_2(t)\right|^2 + c\varepsilon^{\alpha\wedge\mu}. \tag{6.6}$$

Preuve Intéressons nous premièrement à la différence $\bar{X}_1(t) - X_1(t)$.

$$\begin{aligned}
&\mathbf{E}|\bar{X}_1(t) - X_1(t)|^2 \\
\leq\ & 4\,\mathbf{E}\left[\sum_{k=1}^{\infty} \mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{Q_-^k}|\hat{X}_1^-(t) - X_1(t)|^2\right] \\
+\ & 4\,\mathbf{E}\left[\sum_{k=1}^{\infty} \mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{Q_+^k}|\hat{X}_1^+(t) - X_1(t)|^2\right] \\
+\ & 4\,\mathbf{E}\left[\sum_{k=1}^{\infty} \mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a\leq\delta\varepsilon^\mu\}}|X_1(t)|^2\right] \\
+\ & 4\,\mathbf{E}\left[\sum_{k=1}^{\infty} \mathbf{1}_{]\tau_{k-1}^b,\tau_k^a[}(t)|X_1(t)|^2\right].
\end{aligned} \tag{6.7}$$

1ier terme:

$$\begin{aligned}
&\mathbf{E}\left[\sum_{k=1}^{\infty} \mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{Q_-^k}|\hat{X}_1^-(t) - X_1(t)|^2\right] \\
\leq\ & \mathbf{E}\left[\sum_{k=1}^{\infty} \mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{Q_-^k\cap(A_-^k)^c}|\hat{X}_1^-(t) - X_1(t)|^2\right] \\
+\ & \mathbf{E}\left[\sum_{k=1}^{\infty} \mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{A_-^k}|\hat{X}_1^-(t) - X_1(t)|^2\right]. \\
\leq\ & \left[\sum_{k=1}^{\infty}\left(\mathbf{E}\left(\mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\right)\right)^{1/2} P\left(\{\tau_k^b - \tau_k^a > \delta\varepsilon^\mu\}\cap Q_-^k\cap(A_-^k)^c\right)^{1/2}\right]^{1/2} \\
& \qquad \left[\mathbf{E}|\hat{X}_1^-(t) - X_1(t)|^4\right]^{1/2}
\end{aligned}$$

$$+ \quad \frac{1}{|H_1^-|^2}\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{(\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{(\tau_k^b-\tau_k^a>\delta\varepsilon^\mu)}\mathbf{1}_{A_-^k}|H_1^-\hat{X}_1^-(t)-h_1(X_1(t))|^2\right]$$

$$\leq \quad c\varepsilon^{-c/\varepsilon^\beta}$$

$$+ \quad \frac{1}{|H_1^-|^2}\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{(\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{A_-^k}|H_1^-\hat{X}_1^-(t)-h_1(X_1(t))|^2\right],$$

avec $\beta=\frac{1}{2}\wedge(1-\mu)\wedge(\gamma-2\alpha)$, $2\alpha<\gamma<1$, d'après le théorème 5.3 et la finitude des moments de tout ordre de $X_1(t)$ et $\hat{X}_1^-(t)$, pour $t\in[0,T]$.

D'autre part, soit $\theta>0$, en utilisant la proposition 4.1, on obtient:

$$\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{(\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{A_-^k}|H_1^-\hat{X}_1^-(t)-h_1(X_1(t))|^2\right]$$

$$\leq \quad \theta^2\varepsilon^{2\alpha}$$

$$+ \quad P\left(\left\{\sup_{[\varepsilon^\mu,T]}|H_1^-\hat{X}_1^-(t)-h_1(X_1(t))|>\theta\varepsilon^\alpha\right\}\right)^{1/2}$$

$$\left[\mathbf{E}|H_1^-\hat{X}_1^-(t)-h_1(X_1(t))|^4\right]^{1/2}$$

$$\leq \quad c\theta^2\varepsilon^{2\alpha}+c\varepsilon^{-c/\varepsilon^{\gamma-2\alpha}}.$$

$2^{\text{ième}}$ terme: Par un même raisonnement, on montre que

$$\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{(\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a>\delta\varepsilon^\mu\}}\mathbf{1}_{Q_+^k}|\hat{X}_1^+(t)-X_1(t)|^2\right]\leq c\theta^2\varepsilon^{2\alpha}+c\varepsilon^{-c/\varepsilon^\beta}.$$

$3^{\text{ième}}$ terme: D'après la proposition 4.1, il suffit de majorer

$$\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{(\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a\leq\delta\varepsilon^\mu\}}\mathbf{1}_{\left\{\sup_{[\varepsilon^\mu,T]}|H_1^+\hat{X}_1^+(t)-h_1(X_1(t))|\leq\theta\varepsilon^\alpha\right\}}|X_1(t)|^2\right].$$

Or, sur l'événement $\left\{\sup_{[\varepsilon^\mu,T]}|H_1^+\hat{X}_1^+(t)-h_1(X_1(t))|\leq\theta\varepsilon^\alpha\right\}$, on a:

$$|X_1(t)|\leq c|\hat{X}_1^+(t)|+\bar{\theta}\varepsilon^\alpha,\quad\forall t\in[\varepsilon^\mu,T].$$

D'où,

$$\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a\le\delta\epsilon^\mu\}}\mathbf{1}_{\left\{\sup_{[\epsilon^\mu,T]}|H_1^+\hat{X}_1^+(t)-h_1(X_1(t))|\le\theta\epsilon^\alpha\right\}}|X_1(t)|^2\right]$$

$$\le\ \bar{\theta}^2\epsilon^{2\alpha}+c\,\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a\le\delta\epsilon^\mu\}}|\hat{X}_1^+(t)|^2\right].$$

D'après (3.16), $\hat{X}_1^+(t)$ s'écrit

$$\hat{X}_1^+(t) = \quad e^{-\frac{\lambda_1}{\epsilon}(t-\tau_k^a)}c_a\epsilon^\alpha+\frac{K_1}{\epsilon}\int_{\tau_k^a}^t e^{-\frac{\lambda_1}{\epsilon}(t-s)}h(X_s)ds$$

$$+\ \ K_1\int_{\tau_k^a}^t e^{-\frac{\lambda_1}{\epsilon}(t-s)}dW_s,\quad t\in[\tau_k^a,\tau_k^b].$$

En traitant séparément les termes de droite de cette expression, on montre que pour ϵ suffisament petit,

$$\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{[\tau_k^a,\tau_k^b]}(t)\mathbf{1}_{\{\tau_k^b-\tau_k^a\le\delta\epsilon^\mu\}}|\hat{X}_1^+(t)|^2\right]\le c\,\epsilon^{2(\alpha\wedge\mu)}. \tag{6.8}$$

$4^{\text{ième}}$ **terme:** D'après la proposition 4.1, il suffit de s'intéresser à

$$\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{]\tau_{k-1}^b,\tau_k^a[}(t)\mathbf{1}_{\left\{\sup_{[\epsilon^\mu,T]}|H_1^+\hat{X}_1^+(t)-h_1(X_1(t))|\le\theta\epsilon^\alpha\right\}}|X_1(t)|^2\right].$$

Vu que sur l'événement $\left\{\sup_{[\epsilon^\mu,T]}|H_1^+\hat{X}_1^+(t)-h_1(X_1(t))|\le\theta\epsilon^\alpha\right\}$,

$$\begin{aligned}|X_1(t)| &\le c|\hat{X}_1(t)|+\bar{\theta}\epsilon^\alpha\\ &\le c\,c_a\epsilon^\alpha+\bar{\theta}\epsilon^\alpha,\qquad \forall t\in]\tau_{k-1}^b,\tau_k^a[;\end{aligned}$$

on a

$$\mathbf{E}\left[\sum_{k=1}^{\infty}\mathbf{1}_{]\tau_{k-1}^b,\tau_k^a[}(t)\mathbf{1}_{\left\{\sup_{[\epsilon^\mu,T]}|H_1^+\hat{X}_1^+(t)-h_1(X_1(t))|\le\theta\epsilon^\alpha\right\}}|X_1(t)|^2\right]\le c\epsilon^{2\alpha}.$$

Ainsi, la première majoration (6.4) du théorème est démontrée.
La seconde s'obtient par un raisonnement similaire en utilisant l'inégalité

$$|\bar{X}_2(t)-M_2^0(t)|\le|\tilde{X}_2(t)-\tilde{M}_2^0(t)|+|K^0(t)||\bar{X}_1(t)-X_1(t)|.$$

Enfin, par la remarque 2.3 et (6.5), on établit la majoration (6.6).

□

7 Conclusion

Dans ce chapitre, nous avons étudié un problème particulier du filtrage linéaire par morceaux dans le cas où l'erreur de filtrage n'est pas asymptotiquement nulle lorsque le bruit d'observation tend vers zéro. La particularité de ce système réside notamment dans le fait que sa non linéarité dépend uniquement de $X_1(t)$ qui sous "l'hypothèse de détectabilité" $(HD1)$: $|H_1^-| \neq |H_1^+|$, peut être estimée précisement. Un filtre approché a été proposé sur un intervalle de temps $[0,T]$. On a montré que l'erreur d'approximation de la composante $X_1(t)$ était asymptotiquement nulle et que celle de $X_2(t)$ était asymptotiquement optimale au sens où, lorsque $\varepsilon \to 0$, elle tendait vers l'erreur de filtrage du processus $\{X_2(t)\,;\, t \geq 0\}$ observé par $\{X_1(t)\,;\, t \geq 0\}$.

Ces résultats peuvent facilement s'étendre à un système multidimensionel avec $\dim X = n > \dim Y = d$, sous l'hypothèse $\ker H_- = \ker H_+ \subset \Delta_u$, où Δ_u est l'hyperplan séparateur de l'espace $\mathbb{R}^n$.

La généralisation au cas où la partition de $\mathbb{R}^n$ est composée de l polyèdres, $l > 2$, n'est possible que si les composantes correspondantes aux directions de découpage sont "directement" mesurées. Donc, le nombre de directions de découpage distinctes doit être inferieur ou égal à d.

D'autre part, il serait intéressant de considérer des modèles avec petits bruits de dynamique et d'observation sur certaines composantes du processus d'état. On remarque par exemple, que lorsque l'hypothèse $(H4)$ n'est pas vérifiée (cf. [9]) i.e. lorsque la dynamique de la composante mesurée n'est pas bruitée, les filtres de Kalman-Bucy (FK_-) et (FK_+) sont à "mémoires courtes". Cette situation s'apparente à celle traitée dans le chapitre I. On peut déterminer sous deux types "d'hypothèses de détectabilité", $(HD1)$ ou $(HD2)$, des intervalles de linéarité de la fonction h par des tests statistiques. Sur chacun de ces intervalles, la condition initiale de la loi conditionnelle étant sans importance, le filtre de Kalman-Bucy correspondant fournit une "bonne" approximation du filtre optimal.

Le même type d'étude peut être envisagée pour une fonction d'observation injective par morceaux par rapport à la composante "directement" mesurée en exploitant les travaux de W.H. Fleming et E. Pardoux [2].

Bibliographie

[1] W.H. Fleming, D. Ji, E. Pardoux : Piecewise linear filtering with small observation noise, in *Analysis and Optimization of Systems*, Lecture Note in Cont. and Info. Sci. **111**, Springer, 725-729, 1988.

[2] W.H. Fleming, E. Pardoux : Piecewise monotone filtering with small observation noise, *Siam J. Control Optim.* **26**, 261-285, 1989.

[3] U. Haussmann, E. Pardoux : A conditionnally almost linear filtering problem with non Gaussian initial condition, *Stochastics*, **23**, 241-275, 1988.

[4] A.H. Haddad : Linear filtering of singularly perturbed systems, *IEEE Transactions on Automatic Control*, **AC-21**, 515-519, 1976.

[5] I. Karatzas, S.E. Shreve : Brownian Motion and Stochastic Calculus, Springer-Verlag, 1988.

[6] A.J. Krener : The asymptotic approximation of non linear filters by linear filters, *in Proc. 7th MTNS Symposium*, Stockholm, 1985.

[7] H. Kwakernaak, R. Sivan : The Maximally achievable accuracy of linear optimal regulators and linear optimal filters, *IEEE Transactions on Automatic Control*, **AC-16**, 79-86, 1972.

[8] R.CH. Liptzer, A.N. Shyriaev : Statistics of Random Processes Vol.I-II, Springer-Verlag, 1977.

[9] P. Milheiro de Oliveira : Etudes asymptotiques en filtrage non linéaire avec petit bruit d'observation, *Thèse*, Université de Provence, septembre 1990.

[10] P.J. Moylan : A note on Kalma-Bucy filters with zero measurement noise, *IEEE Trans. Inform. Th.* **19**, 263-264, 1974.

[11] J. Picard : Efficiency of the extended Kalman filter for non linear systems with small noise, *Rapport INRIA* , **N1068**, 1989.

[12] S.S. Sachs : Asymptotic analysis of linear filtering problems, *M.S. Thesis*, Departement of Systems Engineering, Case Western Reserve University, Cleveland, Ohio, June 1980.

[13] W.Murray Wonham : Linear Multivariable Control: a Geometric Approach, Springer-Verlag, 1979.

[14] I. Yaesh, B.Z. Bobrovsky, Z. Schuss : Asymptotic analysis of the optimal filtering problem for two dimensional diffusions measured in a low noise channel, *SIAM Journal of Appl. Math.*, **Vol.50**, N4, pp. 1134-1155, 1990.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique.

Résumé:

Dans cette thèse, nous étudions des problèmes de filtrage linéaire par morceaux avec petit bruit d'observation. En exploitant l'hypothèse d'un grand rapport signal sur bruit et la linéarité locale des systèmes considérés, nous proposons, sous certaines "hypothèses de détectabilité", des filtres sous-optimaux facilement calculables construits à l'aide de tests statistiques. Nous obtenons des résultats asymptotiques lorsque le bruit d'observation tend vers zéro.

Mots-clés: Filtrage non linéaire
Systèmes avec petit bruit d'observation
Filtres approchés
Rapport de vraisemblance
Filtre de Kalman-Bucy
Calcul différentiel stochastique

I S B N - 2 - 7261 - 0651 - X

I U . 1 2 8